

Capítulo 9

Teste de hipóteses para proporção e para média

9.1 Princípio de máxima verossimilhança

Um dos pilares de construção de testes de hipóteses é o princípio chamado **máxima verossimilhança**. Ele estará apresentado por meio de sua aplicação à situação descrita no Exemplo 61.

√ **Exemplo 61.** Seu professor pegou um dado equilibrado e uma moeda honesta de R\$ 1, nas cujas faces escreveu 1 e 2, e avisou que pretende escolher um destes objetos, lançar e anunciar o resultado; sua tarefa é adivinhar o objeto lançado.

- ▷ Para atender aos puristas que exigem – com razão – o rigor em definições e discussões, devo completar o enunciado assim: o professor primeiramente lançará uma moeda honesta de 10 centavos para definir se, posteriormente, irá lançar o dado ou a moeda cujas faces tem escrito 1 e 2; digamos, a “cara” na moeda de 10 centavos acarreta o lançamento do dado e a “coroa” o da moeda. É óbvio que o resultado do lançamento da moeda de 10 centavos não será revelado, pois se fosse, saberiam o objeto lançado pelo professor em seguida. Essa moeda auxiliar é necessária para expungir do escopo de nossa análise as estratégias de adivinhação que raciocinam da seguinte maneira: “Acredito que o professor tem preferência pessoal pelo dado, e portanto sempre lançará o dado, e, conseqüentemente, “foi o dado” será sempre a minha resposta.”

Suponha que o número anunciado foi 3. Neste caso, você responderia que, sem dúvida, o objeto lançado era o dado. Certamente, sua resposta adveio do seguinte raciocínio:

- (*) A probabilidade de obter 3 ao lançar a moeda é nula, logo é improvável que a moeda que foi lançada. Por outro lado, a probabilidade de obter 3 ao lançar o dado não é nula. Concluindo, minha resposta é que é o dado que foi lançado.

Suponha agora que o número anunciado foi 2. Neste caso, o raciocínio do tipo de (*) não funciona, pois tanto moeda quanto a dado podem dar 2 com probabilidades não nulas. Mas já que você é obrigado a escolher só um dos objetos, você pensa assim:

- (**) A probabilidade de obter 2 ao lançar a moeda é $1/2$; enquanto que a probabilidade de obter 2 ao lançar o dado é $1/6$. Logo, é mais provável que a moeda tenha sido lançada. Vou alinhar minha adivinhação com aquilo que for mais provável. Concluindo, minha resposta é que é a moeda que foi lançada.

Os pensamentos marcado por (*) e (**), com os quais sua intuição certamente concorda, são exemplos de aplicação daquilo que na Teoria de Probabilidade e Estatística chama-se por princípio de máxima verossimilhança. Meu objetivo agora é expressar (*) e (**) da forma que é a mais próxima à formulação canônica do princípio de máxima verossimilhança. Isso vai lhe permitir uma plena compreensão do princípio.

Para a expressão formal dos raciocínios (*) e (**), é precisa das notações que serão introduzidas abaixo.

O primeiro dos dois conjuntos de conceitos e notações das quais precisaremos, corresponde à visão ao experimento aleatório antes dele acontecer. Especificamente falando, vamos nós posicionar na linha de tempo antes do momento no qual o professor executa seu experimento aleatório. Desse ponto de vista, o número que o professor anunciará é uma variável aleatória. Denotamos essa por X . A partir da descrição do experimento aleatório, possuímos a seguinte informação sobre a distribuição de X :

x	$\mathbb{P}[X = x]$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

A distribuição de X caso for lançado o dado.

x	$\mathbb{P}[X = x]$
1	1/2
2	1/2
3	0
4	0
5	0
6	0

A distribuição de X caso for lançado a moeda está na tabela à esquerda. Essa mesma distribuição está apresentada na tabela à direita da maneira que facilitará a exposição dos futuros argumentos.

Agora vamos nos posicionar após o experimento e vamos definir certos conceitos e notação vistos dessa perspectiva. Denotamos por x_{obs} o número que o professor anunciou. O nome tradicional dele é **observação**, termo que explica as letras “*obs*” usados no seu índice.

- ▷ Observe: x_{obs} denota o que foi observado e anunciado pelo professor, enquanto que x foi usado nas tabelas de distribuição para denotar o valor que pode ser observado.

Recorde que você precisa escolher entre duas alternativas: “foi lançado dado” e “foi lançada moeda”. Estas são chamadas de **hipóteses**. Para serem distinguidas entre si, uma será chamada **hipótese nula** e denotada sucintamente por H , e a outra **hipótese alternativa** com a notação A . No presente exemplo, a atribuição dos nomes é livre e não altera em nada nem o raciocínio nem o resultado da aplicação do mesmo. Seja essa atribuição assim:

H : “foi lançado dado”
 A : “foi lançada moeda”.

- ▷ A regra rígida sobre quem deve ser chamado de “nula” e quem de “alternativa” começa valer – e deve ser obedecida – nos casos quando pelo menos uma das hipóteses não for precisa. Isso será discutido nas seções posteriores.

Com as notações introduzidas, vamos agora á formalização das soluções (*) e (**). Realço que eu não pretendo alterar o raciocínio intuitivo emregado em (*) e (**); só vou as reformular usando termos e notações que são mais genéricos e mais matemáticos.

No caso quando o professor anunciou “deu 3”, tem-se $x_{obs} = 3$ e o argumento (*) adquire a seguinte formalização (as probabilidades abaixo adveiam das tabelas de distribuição de X que foram construídas acima):

(*) $\left. \begin{array}{l} \text{se } H \text{ for verdade, então } \mathbb{P}[X = x_{obs}] = 1/6 \\ \text{se } A \text{ for verdade, então } \mathbb{P}[X = x_{obs}] = 0 \end{array} \right\} \Rightarrow$ Aceitamos aquela das hipóteses que acarreta a maior probabilidade. Como $1/6 > 0$, aceita-se H , quer dizer, conclui-se que foi lançado o dado.

Eu devo acrescentar que a frase “se H for verdade, então $\mathbb{P}[X = x_{obs}] = 1/6$ ” pode ser dita da seguinte maneira: “dado que aconteceu H , tem-se que $\mathbb{P}[X = x_{obs}] = 1/6$ ”. A palavra “dado” da última frase indica que a probabilidade nela contida pode ser vista como probabilidade condicional. A Teoria Estatística adota tal visão e re-escreve toda a frase assim: “ $\mathbb{P}[X = x_{obs} | H] = 1/6$ ”. Eu não gosto dela pois o uso dos símbolos da probabilidade condicional não é mais esclarecedor em comparação com as duas primeiras maneiras, mas confunde as cabeças daqueles alunos que não se sentem firmes com os conceitos ligados à probabilidade condicional. Entretanto, a escrita “cofundadora” existe e está sendo usada amplamente. Eis como aparece o argumento acima com o uso daquela escrita:

$$(*) \left. \begin{array}{l} \mathbb{P}[X = x_{obs} | H] = 1/6 \\ \mathbb{P}[X = x_{obs} | A] = 0 \end{array} \right\} \Rightarrow \begin{array}{l} \text{Aceitamos aquela das hipóteses que} \\ \text{acarreta a maior probabilidade. Como} \\ 1/6 > 0, \text{ aceita-se } H, \text{ quer dizer,} \\ \text{conclui-se que foi lançado o dado.} \end{array}$$

No caso quando o professor anunciou “deu 2”, tem-se $x_{obs} = 2$, o argumento $(**)$ adquire a formalização apresentada abaixo. Recorde que as probabilidades nela usadas adveem das tabelas de distribuição de X construídas acima. Nesta formalização, eu uso a notação de probabilidade condicional com o intuito de acostumar meu leitor com a mesma.

$$(**) \left. \begin{array}{l} \mathbb{P}[X = x_{obs} | H] = 1/6 \\ \mathbb{P}[X = x_{obs} | A] = 1/2 \end{array} \right\} \Rightarrow \begin{array}{l} \text{Aceitamos aquela das hipóteses que} \\ \text{acarreta a maior probabilidade. Já} \\ \text{que } 1/6 < 1/2, \text{ aceita-se } A, \text{ quer dizer,} \\ \text{conclui-se que foi lançada a moeda.} \end{array}$$

A partir das formalizações de nosso raciocínio intuitivo $(*)$, $(**)$ podemos concluir que a intuição seguiu a seguinte regra:

PMV: *escolher aquela das hipóteses, que dá a maior probabilidade ao acontecimento que acarretaria o valor observado.*

Tal regra chama-se **princípio** ou **lei de máxima verossimilhança**. Espero que o presente exemplo exemplificou todos os termos usados na formulação do PMV. Só falta-me acrescentar que o embasamento do PMV é a pura intuição.

Fim do Exemplo 61↑

9.2 Um exemplo com duas hipóteses precisas

- I. Um buffet infantil disponibilizou uma piscina de bolinhas para crianças e contratou uma empresa para fornecer as bolinhas brancas e pretas que preencheriam a piscina. Aconteceu um desentendimento: o administrador do buffet alega que a proporção de bolinhas pretas na piscina é de apenas 0,30, enquanto que o fornecedor das bolinhas acredita que tal proporção é 0,40. O desentendimento afeta o pagamento ao fornecedor, pois as bolinhas pretas são mais caras (para que a cor não desbote com a ação do sol, usa-se um pigmento caro para a coloração).

O desentendimento surgiu devido ao seguinte: o fornecedor contou que dentro das 100 caixas de bolinhas carregadas no caminhão de entrega haviam 40 com bolas pretas, enquanto que ao fiscalizar o descarregamento do caminhão, o administrador do buffet anotou que haviam 30 de tais caixas.

Um dos dois errou na contagem de caixas, mas nenhum dos dois quer recontar todas as bolinhas já que essas foram despejadas das caixas à piscina e misturadas.

Para solucionar o conflito, os dois lados concordaram em proferir o seguinte procedimento

- o escolher aleatoriamente 30 bolinhas, uma após outra, com reposição (quer dizer, devolver cada bolinha escolhida à piscina antes de prosseguir para a próxima escolha); (9.1)

- o caso a quantidade de bolinhas pretas dentro das escolhidas for 10 ou menos, quer dizer, caso tal quantidade for do conjunto de valores $\{0, 1, \dots, 9, 10\}$ tomar-se-á a decisão que são 30% de bolinhas pretas na piscina, mas caso contrário, isto é, caso tal quantidade for 11 ou mais, quer dizer, caso tal quantidade for do conjunto de valores $\{11, 12, \dots, 29, 30\}$ a decisão será que são 40% das bolinhas pretas na piscina. (9.2)

Esse procedimentos será analisado por nos abaixo. Na nossa análise, vamos usar a nomenclatura da Teoria Estatística: o procedimento (9.1)–(9.2) chamar-se-á **procedimento de teste de hipóteses** enquanto que sua parte (9.2) chamar-se-á **regra de decisão**.

- II. A leitura da Parte II é opcional. Nesta parte, eu revelo os motivos que guiaram o administrador e o fornecedor a elaborar a regra de decisão (9.1)–(9.2). Acontece que para você, meu leitor, tais motivos são irrelevantes, pois seu objetivo é aprender a abordagem canônica para a construção de regras de decisão em testes de hipóteses que será lhe apresentada a partir da Parte III da presente seção. É claro que a abordagem canônica baseia-se no bom senso e que por isso que tem pontos em comum com raciocínios de pessoas sensatas, do tipo daquele que foi usado pelos administrador e fornecedor. Mas penetrar em tais raciocínios não lhe ajudará em nada naquilo que tem a ver com sua aprendizagem da abordagem canônica.

- ▷ O administrador e o fornecedor partem para a construção de Regra de Decisão da crença comum de que a proporção amostral é uma boa aproximação para a proporção populacional. Portanto, pensam eles, se obtivermos 9 bolas pretas dentro da amostra de 30 bolas, então a proporção amostral fica em $9/30 = 0,3$, e, conseqüentemente, poderemos decidir que a proporção populacional é 30%. Já se obtivermos 12, então a proporção amostral fica em $12/30 = 0,4$, e poderemos decidir que a proporção populacional é 40%.

Entretanto, com a regra agora construída não podemos simplesmente partir para a amostragem, pois o número de bolinhas pretas na amostra poderá ser diferente de 9 ou 12. O que faremos se observarmos valores diferentes de 9 ou 12? O bom senso sugere que ao observar um número menor que 9 bolinhas pretas na amostra deve-se concluir que a proporção populacional de bolas pretas é 30%, e se observarmos um número maior que 12 deve-se concluir que esta proporção

é 40%. E nesta devo concordar com eles, pois o bom senso comum funciona nesse caso de acordo com o Princípio de Maxima Verossimilhança. O funcionamento puro desse princípio será apresentado na Parte III. No momento só posso dizer que tanto o administrados e o fornecedor quanto o Princípio raciocinam assim: a probabilidade de observar qualquer número menor que 9 é maior sob a hipótese de que a proporção populacional de bolas pretas é 30% do que sob a hipótese de que ela é 40%, e portanto, quando o número de bolas pretas na amostra for menor que 9, devemos concluir que a proporção populacional é 30%. Da maneira similar deduz-se a regra de que o número de bolas pretas na amostra maior que 12 indica que a proporção populacional de bolas pretas na piscina é 40%.

Acontece que sobraram os valores 10 e 11. O que fazer com estes segue-se também do bom senso: dividir igualmente entre as duas afirmações concorrentes usando o princípio intuitivo de proximidade: 10 indicará 30% e 11 indicará 40%. No presente caso, tal divisão parece ser justa para as pessoas envolvidas, e, ainda mais, nossa posterior análise dessa divisão mostrará que a mesma pode ser justificada pelo Princípio de Maxima Verossimilhança. Entretanto, o mesmo bom senso estaria confuso caso o tamanho de amostra fosse 300 e não 30, e caso ele precisasse dividir os valores que ficam entre 90 e 120. A dúvida será que a divisão igualitária por tamanhos pode não significar a divisão igualitária de chances. A dúvida pode ser vista pelas lentes da Análise das Probabilidades de Acertos e Erros que será construída na presnete seção. A sofisticação dessa construção indica que o bom senso não resolveria a dúvida sem forte amparo de teoria.

III. Começaremos agora o análise do procedimento (9.1)–(9.2). Para que você, meu leitor, tenha idéia da meta do análise, imagine como se o administrador do buffet e o fornecedor de bolas desconfiassem da sugestão e viessem para pedir ajuda: “O procedimento é correto? É justo? Podemos aceitá-lo ou devemos procurar por algo melhor?” Saiba, entretanto, que o objetivo real por trás disso é exemplificar os idéias básicas do procedimento de teste de hipóteses.

- ▷ Vale eu avisa-lhe que existem mais que uma maneira de raciocínio que possa ser aplicada à situação acima descrita e resolver a discordância entre duas afirmações. Qualquer um de tais raciocínios pode ser chamado por Teste de Hipóteses para Proporção. Entretanto, nós vamos reservar esse termo somente para o procedimento a ser ensinado no presente capítulo.

É muito importante identificar o parâmetro sobre cujo valor ocorre o desentendimento/discrepância de opiniões. No presente caso ele é a proporção de bolas pretas na piscina. É útil dar a notação para esse parâmetro, pois ele aparecerá em diversos lugares de nossa análise a vir. Então,

$$\text{denotamos por } p \text{ a proporção de bolas pretas na população (piscina)} \quad (9.3)$$

Todo o resto de nossa análise será feita a partir da posição de antes da retirada de amostra. O primeiro passo é identificar a quantia a ser usada na regra de decisão. No presente caso tal quantia é a quantidade de bolas pretas a ser vista na amostra de tamanho 30. Na nossa visão atual, ela é uma variável aleatória. Vamos denotá-la por X . Nesse mesmo passo, tenta-se identificar tudo o possível sobre a distribuição de X . No presente caso, sabemos (a partir de nosso conhecimento das propriedades da Distribuição Binomial) que

$$\begin{aligned} X &\sim \text{Bin}(30, p) \\ \text{isto é, a quantidade de bolas pretas a serem vistas na amostra} & \\ \text{de tamanho 30 tem a distribuição binomial } &\text{Bin}(30; p) \end{aligned} \quad (9.4)$$

- ▷ Seria errado se você pensasse que nosso achado sobre a distribuição de X não serve para nada pelo fato de depender de parâmetro desconhecido p . Muito pelo contrário: o objetivo é expressar tudo que possível em dependência de p . O por quê isso é útil, você vai descobrir dos próximos passos de nossa análise.

O segundo passo da análise do procedimento de teste de hipóteses, identificam-se as hipóteses. Por hipóteses chamamos as afirmações/opiniões confrontantes ou concorrentes da situação analisada. No presente caso, o confronto está claramente entre as opiniões do administrador de buffet e do fornecedor de bolinhas. Uma opinião diz “ $p = 0,3$ ” e outra “ $p = 0,4$ ”. Essa identificação é suficiente para nosso caso, mas já que o objetivo final da presente apresentação é exemplificar o método genérico, então eu vou ter que tratar as opiniões identificadas de acordo com a procedimento canônico de teste de hipóteses, e tal procedimento manda distribuir entre as opiniões os nomes **hipótese nula** e **hipótese alternativa** e usar as notações, respectivamente, H e A ou H_0 e H_a . A distribuição de nomes é importante em geral, mas não para o presente caso. Agora, podemos fazer ela de qualquer maneira. Seja então a distribuição de nomes assim:

$$\begin{aligned} H(\text{a hipótese nula} = \text{a afirmação do administrador}) &: p = 0,3 \\ A(\text{a hipótese alternativa} = \text{a afirmação do fornecedor}) &: p = 0,4 \end{aligned} \tag{9.5}$$

O terceiro passo é descobrir a distribuição da quantia denotada acima por X em duas situações hipotéticas diferentes: uma delas é quando a hipótese nula é válida e a outra – quando a hipótese alternativa. A execução é fácil. No caso da hipótese nula, se assumirmos que ela é correta, então, de acordo com a afirmação da mesma, $p = 0,3$. Isso implica no que X tenha a distribuição $Bin(30; 0,3)$. Os valores dessa distribuição estão apresentados no desenho de cima da Figura 9.1.

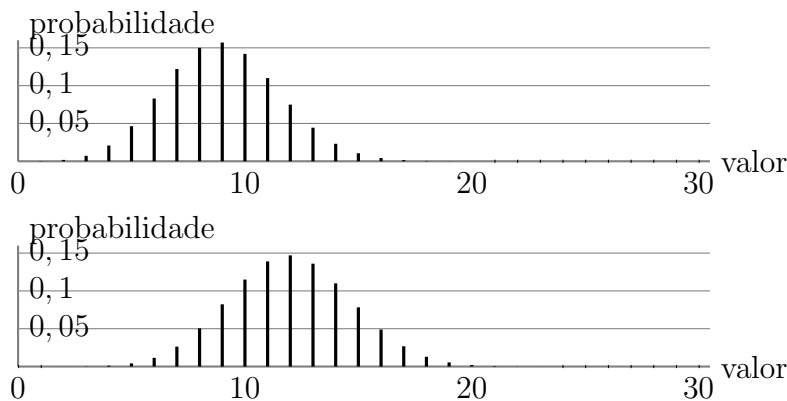


Figura 9.1: Distribuições de $Bin(30; 0,3)$ (acima) e de $Bin(30; 0,4)$ (abaixo). As distribuições estão apresentadas por suas funções-probabilidade: a altura de cada “palito” corresponde à probabilidade de assumir o valor no qual o palito está fincado.

Continuamos ainda no nosso terceiro passo. Agora vamos assumir que a hipótese alternativa está correta. Isso significa que $p = 0,4$, e, conseqüentemente, X tem a distribuição $Bin(30; 0,4)$. Os valores dessa distribuição estão apresentados no desenho de baixo da Figura 9.1.

As duas distribuições foram alinhadas na Figura 9.1 com o intuito de visualizar os seguintes fatos:

- para qualquer $x \in \{0, 1, \dots, 9, 10\}$, tem-se que $P[X = x]$ é maior para $X \sim Bin(30; 0,3)$ de que para $X \sim Bin(30; 0,4)$;
- para qualquer $x \in \{11, 12, \dots, 29, 30\}$, tem-se que $P[X = x]$ é maior para $X \sim Bin(30; 0,4)$ de que para $X \sim Bin(30; 0,3)$

Esses dois fatos podem ser re-escritos na forma apresentada em (9.6) abaixo. A re-escrita expressa por meio da notação de probabilidade condicional os fatos que X tem distribuição

$Bin(30; 0, 3)$ quando a hipótese H for verdade, e tem a distribuição $Bin(30; 0, 4)$ quando a hipótese A for verdade.

$$\begin{aligned} \mathbb{P}[X = x | H] &> \mathbb{P}[X = x | A], \text{ para qualquer } x \in \{0, 1, \dots, 9, 10\}, \\ \mathbb{P}[X = x | H] &< \mathbb{P}[X = x | A], \text{ para qualquer } x \in \{11, 12, \dots, 29, 30\}, \end{aligned}$$

(recordando que H alega que $p = 0, 3$ e acarreta que $X \sim Bin(30; 0, 3)$, enquanto que A alega que $p = 0, 4$ e acarreta que $X \sim Bin(30; 0, 4)$)

(9.6)

Segue-se imediatamente das relações (9.6) – via a aplicação direta do Princípio de Máxima Verossimilhança (explicado na Seção 9.1) – que a escolha entre H e A com base no x_{obs} , o número de bolas pretas na amostra de tamanho 30, deve seguir a seguinte regra: se x_{obs} for um dos valores de $\{0, 1, \dots, 9, 10\}$ deve-se aceitar H , enquanto que se x_{obs} for um dos valores de $\{11, 12, \dots, 29, 30\}$ deve-se aceitar A . Tal regra é exatamente a regra (9.1)–(9.2) acordada entre o fornecedor de bolinhas e o administrador de buffet. Concluindo, podemos responder aos dois que sua regra é boa pois pode ser justificada pelo um dos princípios canonizados da Teoria Estataística.

Mas isto não é o final de minha análise da situação que envolveu o administrador e o fornecedor, pois não é sempre que o Princípio de Máxima Verossimilhança consegue definir a regra de decisão, assim como aconteceu nesta situação. A abordagem genérica emprega também a análise das probabilidades de acertos e erros. Todo isso será explicado em seguida.

IV. Estamos então ainda na análise do exemplo sobre a piscina de bolinhas, e o objetivo atual é mostrar a aplicação do procedimento canônico de Teste de Hipóteses que mistura o Princípio de Máxima Verossimilhança, já apresentado, com a Análise das Probabilidades de Acertos e Erros, que ainda será apresentado.

Na primeira etapa de procedimento, eu aproveito a análise feita acima para concluir que a quantidade de bolas pretas na amostra (que foi denotada pela variável aleatória X) assume valores pequenos com probabilidades maiores quando a hipótese H é válida, e, por outro lado, assume valores grandes com probabilidades maiores quando a hipótese A é válida. Tal conclusão é a leitura “livre” da relação (9.6) deduzida acima rigorosamente. Isto me permite sugerir que seja aceita a hipótese H caso X mostrar valor pequeno, mas que seja aceita A caso for grande. E, por fim, eu sugiro a divisão entre valores pequenos e grandes seja assim: há um inteiro, que está denotado por ℓ e cujo valor será definido em breve, que separa pequenos dos grandes de acordo com o desenho abaixo:

$$\underbrace{\{0, 1, \dots, \ell - 1\}}_{\text{valores pequenos}} \quad \underbrace{\{\ell, \ell + 1, \dots, 30\}}_{\text{valores grandes}} \quad (9.7)$$

Antes de prosseguir com a definição do valor de ℓ , gostaria de chamar a atenção de meu leitor à diferença entre a presente abordagem e a que descrevemos acima na Parte III. Na abordagem acima, comparávamos $\mathbb{P}[X = x | H]$ com $\mathbb{P}[X = x | A]$ para cada valor x entre 0 e 30, e dependendo do resultado de comparação, declavamos se x ia acarretar a aceitação de H ou de A . Tal abordagem era a aplicação direta do Princípio de Máxima Verossimilhança, e volto a ressaltar que o princípio aplicava-se a cada valor de x . Já na presente abordagem, o Princípio de Máxima Verossimilhança está empregado somente para determinar que valores pequenos devem levar à aceitação de H , enquanto que valores grandes devem levar à aceitação de A . Quanto ao limiar ℓ que separa pequenos de grandes, ele será definido pela Análise das Probabilidades de Acertos e Erros o que faremos a partir do parágrafo seguinte.

Atenção! Neste exato momento entra em cena a ideia em cima da qual será construído a Análise das Probabilidades de Acertos e Erros. A idéia é que a Regra de Decisão pode acertar

	A Regra de Decisão decide que	
	a hipótese H é verdade	a hipótese A é verdade
A verdade está com a hipótese H	acertou	errou
A verdade está com a hipótese A	errou	acertou

Figura 9.2: A tabela de todos os possíveis acertos e erros de Regra de Decisão.

e errar, e que são um erro e um acerto para cada realidade virtual correspondente a cada uma de duas hipóteses. A tabela da Figura 9.2 esclarece por definitivo essa idéia.

Vamos ao cálculo das probabilidades de acertos e erros definidos pela tabela da Figura 9.2. Faremos isso para diversos valores de ℓ , e depois, ao analisar as probabilidades, escolheremos o melhor valor para ℓ . (O critério para “o melhor” estará formulado como “O motivo para aceitar $\mathcal{RC} = \{11, 12, \dots, 30\}$ ” que será formulado perto do final da presnete Parte IV.)

Os cálculos a ser apresentados abaixo incorporarão em suas passagens um novo conceito. É tal de “Região Crítica”. Este conceito é inseparável de toda a teoria de Teste de Hipóteses. Por isto, quanto mais cedo eu introduzo ele na minha apresentação, mais tempo terá meu leitor para se acostumar com ele. **Região Crítica** é o nome para o conjunto de valores para os quais a hipótese nula será rejeitada. Por exemplo, se no acordo (9.1)–(9.2) entre o fornecedor e o administrador tomarmos a afirmação do administrador (que é que $p = 0, 3$) como a hipótese nula, então a correspondente Região Crítica será $\{11, 12, \dots, 29, 30\}$. Vale observar que a Região Crítica vincula-se à hipótese nula. Tanto é que se, por exemplo, no mesmo acordo (9.1)–(9.2) a gente chamar de hipótese nula a afirmação do fornecedor (que é que $p = 0, 4$) então a Região Crítica torna a ser $\{0, 1, \dots, 9, 10\}$. Tal camaleaoicidade incomoda e muito. Por mim, o termo alternativo melhor seria “a região de rejeção da hipótese nula”. Mas este, infelizmente, não é o termo padrão na Teoria Estatística.

Nosso primeiro cálculo das probabilidades para acertos e erros será feito agora para $\ell = 11$, quer dizer para o caso no qual a região de rejeição da hipótese nula (que é que $p = 0, 3$) é $\mathcal{RC} = \{11, 12, \dots, 29, 30\}$.

- (a) Se H for a verdade, então $X \sim Bin(30; 0, 3)$, e a Regra de Decisão acerta em H caso o valor de X estar no conjunto $\{0, 1, \dots, 9, 10\}$, o que acontece com a probabilidade

$$\mathbb{P}[X \in \{0, 1, \dots, 9, 10\}], \text{ onde } X \sim Bin(30; 0, 3), \quad (9.8)$$

cujo valor numérico é (com o uso da Tabela da Distribuição de $Bin(30; 0, 3)$ apresentada na Figura 9.3) 0, 73.

- (b) Se H for a verdade, então $X \sim Bin(30; 0, 3)$, e a Regra de Decisão erra, ao apontar em A como a verdade, caso o valor de X estar no conjunto $\{11, 12, \dots, 29, 30\}$, o que acontece com a probabilidade

$$\mathbb{P}[X \in \{11, 12, \dots, 29, 30\}], \text{ onde } X \sim Bin(30; 0, 3), \quad (9.9)$$

cujo valor numérico é (com o uso da Tabela da Distribuição de $Bin(30; 0, 3)$ apresentada na Figura 9.3) 0, 27.

k	$\mathbb{P}[X = k]$ para $X \sim \text{Bin}(30; 0, 3)$	$\mathbb{P}[X = k]$ para $X \sim \text{Bin}(30; 0, 4)$	$\mathbb{P}[X = k]$ para $X \sim \text{Bin}(20; 0, 3)$	$\mathbb{P}[X = k]$ para $X \sim \text{Bin}(20; 0, 4)$
0	0.000023	0.00000022	0.00079792266297612	0.0000365615844
1	0.0003	0.0000044	0.00683933711122388	0.0004874877920
2	0.0018	0.0000427	0.02784587252426865	0.0030874226827
3	0.0072	0.000265	0.07160367220526231	0.0123496907308
4	0.0208	0.0012	0.13042097437387065	0.0349907904041
5	0.0464	0.00415	0.17886305056987975	0.0746470195288
6	0.083	0.0115	0.19163898275344257	0.1244116992147
7	0.122	0.0263	0.16426198521723651	0.1658822656197
8	0.150	0.0505	0.11439673970486122	0.1797057877546
9	0.157	0.0823	0.06536956554563482	0.1597384780041
10	0.142	0.115	0.03081708090008504	0.1171415505363
11	0.110	0.139	0.01200665489613703	0.0709948791129
12	0.0749	0.147	0.00385928193090119	0.0354974395564
13	0.0444	0.136	0.00101783259716075	0.0145630521257
14	0.02312	0.110	0.00021810698510587	0.0048543507085
15	0.0106	0.0783	0.00003738976887529	0.0012944935222
16	0.00424	0.0489	0.00000500755833151	0.0002696861504
17	0.0015	0.0268	0.00000050496386536	0.0000423037098
18	0.00046	0.0129	0.00000003606884753	0.0000047004122
19	0.000125	0.00548	0.00000000162716605	0.0000003298534
20	0.00003	0.001997	0.00000000003486784	0.0000000109951
21	0.000006	0.000634		
22	0.000001	0.000173		
23	0.00000016	0.000040		
24	0.0+	0.0+		
25	0.0+	0.0+		
26	0.0+	0.0+		
27	0.0+	0.0+		
28	0.0+	0.0+		
29	0.0+	0.0+		
30	0.0+	0.0+		

Figura 9.3: As distribuições binomiais usadas nos cálculos de probabilidades de acertos e erros da Seção 9.2. As distribuições são necessárias também para resolução dos Exc. 142–145 e 148–150. A notação “0.0+” significa que o correspondente valor é estritamente maior que 0 mas que todos os algarismos anteriores ao da casa decimal que queremos arredondar foram igual a 0.

- (c) Se A for a verdade, então $X \sim \text{Bin}(30; 0, 4)$, e a Regra de Decisão acerta em A caso o valor de X estar no conjunto $\{11, 12, \dots, 29, 30\}$, o que acontece com a probabilidade

$$\mathbb{P}[X \in \{11, 12, \dots, 29, 30\}], \text{ onde } X \sim \text{Bin}(30; 0, 4), \quad (9.10)$$

cujo valor numérico é (com o uso da Tabela da Distribuição de $\text{Bin}(30; 0, 4)$ apresentada na Figura 9.3) 0,71.

- (c) Se A for a verdade, então $X \sim \text{Bin}(30; 0, 4)$, e a Regra de Decisão erra, ao apontar em H como a verdade, caso o valor de X estar no conjunto $\{0, 1, \dots, 9, 10\}$, o que acontece com a probabilidade

$$\mathbb{P}[X \in \{0, 1, \dots, 9, 10\}], \text{ onde } X \sim \text{Bin}(30; 0, 4), \quad (9.11)$$

cujo valor numérico é (com o uso da Tabela da Distribuição de $\text{Bin}(30; 0, 4)$ apresentada na Figura 9.3) 0,29.

Observação: Os quem gosta de usar a notação de probabilidade condicional, escreveria as expressões (9.8), (9.9), (9.10) e (9.11) como, respectivamente, $\mathbb{P}[X \in \{0, 1, \dots, 9, 10\} | H]$, $\mathbb{P}[X \in \{11, 12, \dots, 29, 30\} | H]$, $\mathbb{P}[X \in \{11, 12, \dots, 29, 30\} | A]$ e $\mathbb{P}[X \in \{0, 1, \dots, 9, 10\} | A]$.

Em seguida, vamos tornar nossa atenção às probabilidades para acertos e erros no caso $\ell = 10$, quer dizer no caso quando a região de rejeição da hipótese nula (que é que $p = 0, 3$) é $\mathcal{RC} = \{10, 11, 12, \dots, 29, 30\}$. Fica óbvio que os argumentos (a)-(c) aplicadas acima ao caso $\ell = 11$ aplicam-se também para o caso $\ell = 10$ mas com as seguintes alterações dos conjuntos:

$$\{0, 1, \dots, 9, 10\} \text{ para } \{0, 1, \dots, 9\} \text{ e } \{11, 12, \dots, 29, 30\} \text{ para } \{10, 11, \dots, 29, 30\}$$

Em particular, quando H é verdade, a probabilidade de acerto diminui-se (pois na expressão (9.8), o conjunto $\{0, 1, \dots, 9, 10\}$ está substituído pelo conjunto menor $\{0, 1, \dots, 9\}$) e a probabilidade de erro aumenta-se (pois na expressão (9.9), o conjunto $\{11, 12, \dots, 29, 30\}$ está substituído pelo conjunto maior $\{10, 11, \dots, 29, 30\}$). Da maneira similar, deduz-se que quando A é verdade, a probabilidade de acerto aumenta-se e a probabilidade de erro diminui-se.

O argumento apresentado no parágrafo acima mostra que as mesmas tendências de aumento e diminuição se preservam se a gente diminui ℓ para baixo de 9 (veja a tabela na Figura 9.4).

Agora, diferentemente daquilo que fizemos em dois parágrafos acima onde diminuimos ℓ , vamos aumenar ℓ a partir de seu valor original 11 para 12. Em consequência deste aumento, a região de rejeição da hipótese nula (a que alega que $p = 0, 3$) torna-se $\mathcal{RC} = \{12, \dots, 29, 30\}$. Vamos analisar a alteração das probabilidades de acertos e erros. Assim como no caso considerado anteriormente, fica claro que os argumentos (a)-(c) aplicam-se para o caso atual $\ell = 12$, mas que exigem as seguintes alterações de conjuntos

$$\{0, 1, \dots, 9, 10\} \text{ para } \{0, 1, \dots, 10, 11\} \text{ e } \{11, 12, \dots, 29, 30\} \text{ para } \{12, 13, \dots, 29, 30\}$$

Tais alterações acarretas as seguintes mudanças: quando H é verdade, a probabilidade de acerto aumenta-se (pois na expressão (9.8), o conjunto $\{0, 1, \dots, 9, 10\}$ está substituído pelo conjunto maior $\{0, 1, \dots, 10, 11\}$) e a probabilidade de erro diminui-se (pois na expressão (9.9), o conjunto $\{11, 12, \dots, 29, 30\}$ está substituído pelo conjunto menor $\{12, 13, \dots, 29, 30\}$). Da maneira similar, deduz-se que quando A é verdade, a probabilidade de acerto diminui-se e a probabilidade de erro aumenta-se.

	A probabilidade da Regra de Decisão decidir que		
	a hipótese H é verdade	a hipótese A é verdade	para a Região Crítica
A verdade está com a hipótese H	aumenta-se	diminui-se	diminui-se para $\{30\}$
	0,84	0,16	$\mathcal{RC} = \{12, 13, \dots, 30\}$
	0,73	0,27	$\mathcal{RC} = \{11, 12, \dots, 30\}$
	0,59	0,41	$\mathcal{RC} = \{10, 11, \dots, 30\}$
	diminui-se	aumenta-se	aumenta-se para $\{0, \dots, 30\}$
A verdade está com a hipótese A	aumenta-se	diminui-se	diminui-se para $\{30\}$
	0,57	0,43	$\mathcal{RC} = \{12, 13, \dots, 30\}$
	0,29	0,71	$\mathcal{RC} = \{11, 12, \dots, 30\}$
	0,17	0,83	$\mathcal{RC} = \{10, 11, \dots, 30\}$
	diminui-se	aumenta-se	aumenta-se para $\{0, \dots, 30\}$

Figura 9.4: A tabela das probabilidades de acertos e erros.

O argumento apresentado no parágrafo acima mostra que as mesmas tendências de aumento e diminuição se preservam se a gente aumenta ℓ para cima de 11 (veja a tabela na Figura 9.4).

A tabela das probabilidades de acertos e erros mostra que a Regra de Decisão mais equilibrada ocorre quando a Região Crítica for $\{11, 12, \dots, 30\}$, isto é, quando, com a quantidade de bolinhas pretas na amostra entre 0 e 10, a Regra de Decisão decide que H é verdade, e, com a quantidade de bolinhas entre 11 e 30, a Regra de Decisão decide que A é verdade.

Eis o argumento porquê o administrador deve concordar com a Região Crítica $\{11, 12, \dots, 30\}$.

O motivo para aceitar a $\mathcal{RC} = \{11, 12, \dots, 30\}$: Você está convencido que a verdade está a seu lado, quer dizer que a proporção de bolas pretas na piscina, p , é 0,3, e portanto, você deve não gostar nenhum pouco do fato que a regra de decisão pode decidir $p = 0,4$ embora o verdadeiro valor de p é 0,3. Porém, entenda que o fornecedor pensa semelhantemente e também não gosta que a Regra de Decisão pode apontar no que $p = 0,3$ embora, de acordo com a convenção dele, $p = 0,4$. Entretanto, veja que com a $\mathcal{RC} = \{11, 12, \dots, 30\}$, os erros, vistos por você e pelo fornecedor, têm probabilidades quase iguais (conforme mostra Figura 9.4), igualdade que não acontece para nenhuma outra Região Crítica (conforme mostra a mesma figura).

É claro que por um raciocínio similar podemos convencer também o fornecedor de que, para resolver seu desentendimento com o administrador, ele deve aceitar a Regra de Decisão com a Região Crítica $\{11, 12, \dots, 30\}$.

Esta foi nossa justificativa para a regra (9.1)–(9.2). Recorde que amparar a regra era nosso objetivo. O outro objetivo era exibir como isto pode ser feito via a **Análise das Probabilidades de Acertos e Erros**. Espero que a exibição também esclareceu plenamente a essência desta análise. Para aqueles que ainda têm dúvidas a respeito, posso dizer que a análise constitui-se basicamente de duas etapas: identificar os caminhos pelos quais surgem acertos e erros (no caso do presente exemplo, o resultado de tal identificação foi a tabela da Figura 9.2), e calcular as probabilidades de tais acertos e erros (no caso do presente exemplo, o resultado de tal cálculo foi a tabela da Figura 9.4)

V. Vamos agora imaginar que o administrador e o fornecedor foram executar o procedimento de teste e que ao fazer amostra, contaram nela 10 bolas pretas. Vamos ver como fica o desfecho de toda a história.

A regra decide: A proporção populacional de bolas pretas na piscina é 0,3; isto é, $p = 0,3$. (Observe que a conclusão “a proporção populacional = 0,3” foi feita não pelo fato da proporção amostral ser $10/30 = 0,3$, mas de acordo com a Regra de Decisão (9.1)–(9.2).)

O administrador fala para seu chefe: “Minha crença de que p é 30% foi confirmada. Vamos pagar para o fornecedor de acordo com a conta feita na base de crença que a proporção de bolas pretas na piscina é 30%.”

O fornecedor fala para seu chefe: “Foi decidido que $p = 0,3$. Entretanto, isso não garante que p de fato é 0,3. A decisão foi emitida via execução de uma regra, e a mesma possa errar nos dois sentidos, conforme foi mostrado pelo estatístico na sua análise da regra. Em outras palavras, eu tenho direito de continuar com a minha crença de que a quantidade de bolas pretas na piscina é 40%. É que mesmo com $p = 40\%$ há probabilidade de termos 10 ou menos bolas pretas na amostra de tamanho 30. Por mim, é isso que aconteceu. Mas, infelizmente, temos que honrar o acordo com o administrador, e vamos receber dinheiro de acordo a conta feita na base de crença que a proporção de bolas pretas na piscina é 30%.”

O que um estatístico comentaria se fosse perguntado: “Honestamente falando, ninguém sabe por certo se p (a proporção de bolas pretas na piscina) é 30% ou 40%. É que deu 10 bolas pretas na amostra de 30 bolas retiradas da piscina, e, de acordo com a Regra de Decisão fixada antemão, decidimos que p é 30%. Esses 10 bolas pretas em 30 retiradas poderiam aparecer tanto no caso $p = 30\%$ quanto no caso $p = 40\%$.”

VI. Antes de encaminhar meu leitor às próximas seções do presente capítulo, eu gostaria de sugerir que estejam feitos os exercícios da Seção 9.11.1. Tais exercícios foram feitos para fixar as ideias e métodos explicados na presente seção. Isto é importante para que possamos prosseguir para as seções seguintes nas quais discutiremos as situações mais genéricas e mais próximas à realidade de que a situação sobre piscina de bolinhas discutida agora. Com o intuito de facilitar a compreensão das situações a serem discutidas, eu introduzi nos exercícios da Seção 9.11.1 o conceito de nível de significância, pois o mesmo desempenha um papel importante na teoria geral de teste de hipóteses. Eis a definição:

O nível de significância é a probabilidade de rejeitar a hipótese nula quando ela é verdade.

A notação tradicional para o Nível de Significância é α .

Vale notar que o nível de significância é a probabilidade de um dos erros que foram identificados na nossa análise da situação com piscina de bolinhas. Como os erros são dois mas o nível de significância é um só, é preciso entender bem quem dos dois vai ganhar esse nome. Acontece que isso depende de nossa atribuição do nome “hipótese nula”. Por exemplo, no tratamento da situação de piscina de bolinhas, a afirmação “ $p = 0,3$ ” foi escolhida para ser chamada de hipótese nula. Neste caso, o nível de significância é a probabilidade de rejeitar que a afirmação $p = 0,3$ enquanto que na verdade ela é verdade. O valor do nível de significância é 0,27 quando a $\mathcal{RC} = \{11, 12, \dots, 30\}$, conforme mostra a tabela da Figura 9.4. Se chamássemos a afirmação “ $p = 0,4$ ” pela “hipótese nula”, o nível de significância seria a probabilidade de rejeitar que a afirmação $p = 0,4$ enquanto que na verdade ela é verdade.

Observação: O valor de Nível de Significância muda com mudança da Região Crítica. Isso fica claro a partir das definições desses conceitos e da informação apresentada pela tabela da Figura 9.4. Acontece que a amarração entre os dois é a fonte da maioria das perguntas nos exercícios sobre o tema Teste de Hipóteses. Especificamente falando, a maioria das perguntas pedem calcular o nível de significância para uma dada região crítica, ou pedem a recíproca, dado o desejado valor do nível de significância pedem achar a região crítica que acarreta tal valor.

9.3 Teoria geral

A situação “Piscina de bolinhas” é um exemplo de problema que foi resolvido pelo método chamado “teste de hipóteses”. No exemplo, ambas as hipóteses eram precisas, quer dizer cada uma alegava que a proporção populacional de bolas pretas na piscina era um valor específico. Este não é o único caso a ser considerado no presente texto. Há mais dois casos: um é quando só uma hipótese é precisa, e outro quando ambas as hipóteses são imprecisas.

Os problemas dos três tipos descritos acima serão tratados pelo **método de teste de hipóteses**. A presente seção especifica as etapas deste método.

Na **primeira etapa**, você deve identificar que o cerne do problema é escolher uma entre duas alternativas. A presença de duas alternativas concorrentes (ou confrontantes) é a característica principal dos problemas de teste de hipóteses.

- ▷ Se não houverem alternativas confrontantes, o problema admite abordagem por outros métodos estatísticos e surge a questão de comparação dos métodos aplicáveis. Por outro lado, problemas sem alternativas confrontantes podem ser alterados por introdução artificial de alternativa confrontante àquela que representa a questão na formulação original. Nenhum destes dois casos surgirá em nossos exercícios ao longo do curso.

As alternativas são chamadas **hipóteses**. Até um certo momento do texto, vou me referir a estas como 1-a e 2-a hipótese. Porém mais tarde usaremos a nomenclatura tradicional: **hipótese nula** e **hipótese alternativa**.

Na **segunda etapa**, você deve identificar a quantia que será observada e que será usada para escolher qual das duas hipóteses é válida e qual não. Por exemplo, na situação “piscina de bolinhas”, tal quantia era a quantidade de bolas pretas a ser vista na amostra de bolas retiradas da piscina.

- ▷ Em todos os exercícios do presente tema, eu vou descrever – com a devida precisão e cuidado – a amostragem que será feita e a quantia que será observada nesta amostragem. Em caso reais, o método de amostragem e a quantia a ser observada na amostragem são sugeridos pelo estatístico. Este trabalho não será sua tarefa em nenhum exercício do curso.

Na **terceira etapa**, você deve analisar a quantia observada com o objetivo de identificar sua distribuição. Preste atenção: você deve fazer esta análise duas vezes, ou seja, uma vez assumindo que a 1-a hipótese é verdadeira, e a outra vez assumindo que a 2-a hipótese é verdadeira. Por exemplo, na situação “piscina de bolinhas”, esta análise revelou que a quantidade de bolas pretas na amostra tem a distribuição $\text{Bin}(30; 0,3)$ caso a hipótese assumida como a verdadeira for “a proporção de bolas pretas na piscina é 0,3”, e, por outro lado, a distribuição da quantidade de bolas na amostra é $\text{Bin}(30; 0,4)$ caso a hipótese assumida como a verdadeira for “a proporção de bolas pretas na piscina é 0,4”.

- ▷ **Preste atenção:** assumir que uma hipótese seja verdadeira (para o fim de fazer certas contas) não significa aceitar tal hipótese como a verdade. A aceitação/rejeição de hipóteses é uma das últimas etapas do método.

Na **quarta etapa**, você separa os valores possíveis da quantia observada em dois conjuntos de acordo com a seguinte lógica:

- (a) se o valor da quantia observada cair no 1° conjunto, terá que ser aceita como verdadeira a 1-a hipótese;

!ht

	A regra de decisão decide que	
	hipótese 1-a é verdade	hipótese 2-a é verdade
A verdade está com a hipótese 1-a	acertou	errou
A verdade está com a hipótese 2-a	errou	acertou

Figura 9.5:

(b) se o valor da quantia observada cair no 2º conjunto, terá que ser aceita como verdadeira a 2-a hipótese.

Por exemplo, na situação “piscina de bolinhas”, foi decidido que

- caso a quantidade de bolinhas pretas na amostra for 10 ou menos, tomar-se-á a decisão que a proporção de bolinhas pretas na piscina é 0,3;
- caso a quantidade de bolinhas pretas na amostra for 11 ou mais, a decisão será que são a proporção de bolinhas pretas na piscina é 0,4.

Consequentemente, nesta situação, os conjuntos de separação são $\{0, 1, \dots, 10\}$ e $\{11, 12, \dots, 30\}$.

Em geral, a execução desta etapa não é nada trivial. É por isto que nos exercícios do tema, eu vou ou descrever tais conjuntos explicitamente, ou vou fornecer dicas que são suficientes para que você não tenha dúvidas sobre a determinação dos conjuntos.

Ao despeito de você receber ajuda completa para a determinação dos conjuntos de separação, é útil você aprender a regra geral para a construção de tais conjuntos. A grosso modo, o motor da regra funciona assim: ele olha em cada valor que a quantia observada pode assumir e e colocá-lo no conjunto “aceitar-se-á a 1-a hipótese” ou “aceitar-se-á a 2-a hipótese” de acordo com o princípio de máxima verossimilhança. Recerdoe que este princípio funciona na base de comparação de probabilidades. É por isto que na etapa anterior, na terceira quer dizer, você deveria ter identificado as distribuições da quantia observada. Acontece que o princípio de máxima verossimilhança não sempre funciona bem, fato que me motivou chamar a presente etapa de difícil e, conseqüentemente, acrscentar a ajuda adequada a todos os exercícios do tema.

Ainda devo notar que a execução da quarta etapa está ligada aos seguintes termos tradicionais do método de teste de hipóteses: **o procedimento de teste** ou **regra de decisão**; estes são os nome para as instruções (a)-(b) fromadas na quarta etapa.

A **quinta etapa** parece como uma continuação da terceira. Nesta, tem-se duas tarefas:

- você assume que a hipótese 1-a seja verdadeira e calcula a probabilidade que a regra de decisão leve ao fato de que esta hipótese é verdadeira, e calcula também a probabilidade que a regra decida que esta hipótese é falsa;
- você assume que a hipótese 2-a seja verdadeira e calcula a probabilidade que a regra de decisão leve ao fato de que esta hipótese é verdadeira, e calcula também a probabilidade que a regra decida que esta hipótese é falsa.

A tabela da Figura 9.5 apresenta todos os erros e acertos cujas probabilidades devem ser calculadas na quinta etapa.

!ht

	A regra de decisão decide que	
	a hipótese nula é verdade	a hipótese alternativa é verdade
A verdade está com a hipótese nula	acerto sem nome	erro tipo I
A verdade está com a hipótese alternativa	erro tipo II	acerto sem nome

Figura 9.6:

As probabilidades calculadas na quinta etapa possuem nomes específicos na Teoria Estatística. Tais nomes aparecem inevitavelmente nas conclusões de aplicação de Teste de Hipóteses. Isto torna necessária a introdução dos nomes. Tal introdução, porém, carrega com ela um punhado de outros nomes e conceitos. Tudo isto está abaixo.

Recordo em primeiro lugar, que na Teoria Estatística, uma das hipóteses chama-se **nula** e outra chama-se **alternativa**. As correspondentes notações são assim:

- ao invés de dizer “a hipótese nula” escreve-se H_0 ou H ;
- ao invés de dizer “a hipótese alternativa” escreve-se H_A ou A ;
- então observe que dois pares de notações: $H_0 - H_A$ e $H - A$.

Aos nomes “nula” e “alternativa” estão vinculadas as definições de **erro tipo I** e **erro tipo II**. As definições estão abaixo, mas vale avisar que a memorização delas fica mais fácil se for usada a esquema apresentada na tabela da Figura 9.6.

Erro tipo I ocorre quando a hipótese nula é verdadeira, mas a regra de decisão decide que ela é falsa.

Erro tipo II ocorre quando a hipótese alternativa é verdadeira, mas a regra de decisão decide que ela é falsa.

Os erros de dois tipos entram na definição dos conceitos “nível de significância” e “poder de teste” que são os conceitos que acompanham a decisão de teste de hipóteses. A definição está abaixo, e o uso deles estará mostrado na descrição da sexta etapa.

- A probabilidade do erro tipo I é chamado o **nível de significância** do teste. Este nível é denotado por α .
- A probabilidade do erro tipo II denota-se por β ; e $1 - \beta$ chama-se o **poder do teste**.

Para a completude, devo ainda informá-lo que a região dos valores da quantia observada que faz a regra de decisão decidir contra a hipótese nula chama-se **região crítica**. Este é o nome para um dos conjuntos sobre os quais falávamos na quarta etapa. Naquele instante não revelamos seu nome pois o mesmo está vinculado à escolha de nomes para as hipóteses, que não tinha sido discutido até a quarta etapa.

A **sexta etapa** destaca-se das demais pois para sua execução é necessário ter o resultado da amostra. Então, suponhamos que tenham este, e denotamos ele por k . Por exemplo, na situação “piscina de bolinhas”, este k é o número de bolas pretas vistas na amostra. Com o

valor numérico de k , executa-se a Regra de Decisão (formulada na quarta etapa). Esta execução é nada mais e nada menos de que a simples verificação se k pertence à Região Crítica ou não. Com o resultado, será aceita ou a hipótese nula ou a hipótese alternativa. A aceitação/rejeição é o que se informa ao interessado que encomendou o teste. Mas junto com a resposta informam-se os valores do nível de significância e do poder. Estes valores permitem ao estatístico avaliar o teste do ponto de vista de suas características que podem ser chamadas de “sensibilidade”, “eficiência” e “viés em relação a cada hipótese”. Ensinar os princípios desta interpretação é um dos objetivos da apresentação deste capítulo.

√ **Exemplo 62.** Vamos ver aqui a aplicação de tudo ensinado na presente seção à situação de Piscina de Bolinhas assumindo que na amostra de 30 bolas retiradas da piscina observou-se 10 bolas pretas.

Se chamarmos:

- $p = 0,4$ por hipótese **nula** (H),
- $p = 0,4$ por hipótese **alternativa** (A),

e se escolhermos a **região crítica** $\{0, 1, \dots, 9, 10\}$ para

a **Regra de Decisão**: decidir que H é verdade se k (o número de bolas pretas de amostra) estar fora da região crítica, e decidir que A é verdade, caso contrário,

então o **nível de significância** será de 0,29, e o **poder** será de 0,73.

E com $k = 10$, tem-se que k está dentro da Região Crítica, e, portanto, a resposta do teste será: “a teste de hipótese $H : p = 0,4$ contra $A : p = 0,3$, cujo nível de significância é 0,29 e cujo poder é 0,73, rejeitou a hipótese nula”.

Se chamarmos:

- $p = 0,3$ por hipótese **nula** (H),
- $p = 0,4$ por hipótese **alternativa** (A),

e se escolhermos a **região crítica** $\{11, 12, 13, \dots, 29, 30\}$ para a mesma **Regra de Decisão**, então o **nível de significância** será de 0,27, e o **poder** será de 0,71.

E com $k = 10$, tem-se que k está fora da Região Crítica, e, portanto, a resposta do teste será: “a teste de hipótese $H : p = 0,3$ contra $A : p = 0,4$, cujo nível de significância é 0,27 e cujo poder é 0,71, aceitou a hipótese nula.

É muito importante observar que no caso de duas hipóteses precisas, a troca de nomes (nula/alternativa) causa uma reviravolta de outros nomes, mas a decisão final não muda.

_____↑

9.4 Só uma das hipóteses é precisa

Quando deseja-se aplicar o método de teste de hipóteses para escolher entre duas hipóteses na situação em qual uma delas é precisa e outra não, a regra é dar o nome “nula” para aquela das duas que é precisa. Esta não é uma regra absoluta, mas ela deve ser obedecida no âmbito do presente texto.

↓ **Exemplo 63** (como verificar se moeda é honesta). Recebemos uma moeda. Queremos saber se esta é honesta ou não, ou seja, queremos saber se ao ser lançada, a probabilidade da moeda dar “cara” é 0,5 ou não. Podemos lançar esta moeda 10 vezes. (Este 10 é só para fazer com que o tamanho de amostra esteja definido por natureza do problema, caso que acontece frequentemente na prática.) O objetivo do exemplo é construir um teste que possa responder, com base no resultado dos 10 lançamentos, se a moeda é honesta ou não.

Vamos abordar este problema pelo método de teste de hipóteses.

Primeira etapa: Observe, que temos duas alternativas entre as quais devemos escolher:

a moeda é honesta, e

a moeda é desonesta.

Na teoria sobre o teste de hipóteses, estas chamam-se **hipóteses**. Então temos:

1-a hipótese: a moeda é honesta, e

2-a hipótese: a moeda é desonesta.

Usando a notação p para a probabilidade da moeda dar cara num lançamento, podemos reescrever assim:

1-a hipótese: $p = 0,5$ e

2-a hipótese: $p \neq 0,5$, ou em termos afirmativos, $p \in (0,1) \setminus \{0,5\}$, quer dizer, qualquer número no intervalo $(0,1)$, menos $0,5$.

A última maneira de escrever as hipóteses deixa clara o significado do termo **hipótese precisa**. No caso, a 1-a hipótese é precisa enquanto que a 2-a não é. Nos casos deste tipo de problema, reina a seguinte regra: aquela das hipóteses que é precisa recebe o nome NULA, e a outra recebe o nome ALTERNATIVA.

Então, no nosso exemplo, temos:

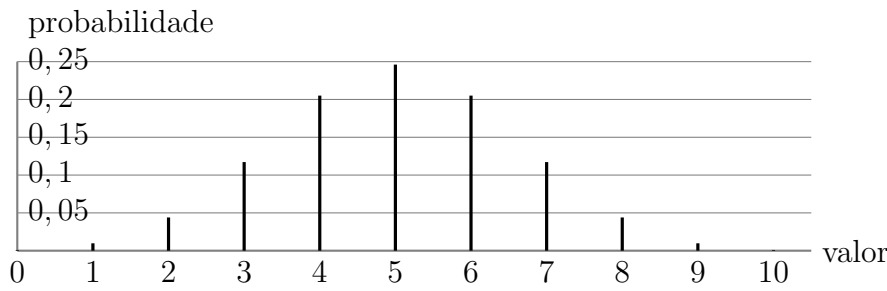
hipótese nula (H): $p = 0,5$ e

hipótese alternativa (A): $p \neq 0,5$.

Na **segunda etapa** da solução, você precisa determinar a quantia a ser observada. O enunciado disse que a moeda pode ser lançada 10 vezes. É natural observarmos o número de caras nestes lançamentos e tomarmos a decisão (acerca de qual das duas hipóteses é verdade) com base no valor desta quantia.

Na **terceira etapa** da solução você precisa achar a distribuição da quantia observada; você precisa fazer isto duas vezes: uma vez assumindo que H é verdade, e a outra, assumindo que A é verdade.

Assumimos que H é verdade. Isto é, assumimos que $p = 0,5$ (p foi introduzido como a notação para a probabilidade da moeda dar “cara” num lançamento). Neste caso, a distribuição do número de “caras” a serem vistas em 10 lançamentos é Binomial(10; 0,5). Veja o desenho desta distribuição na transparência a seguir.



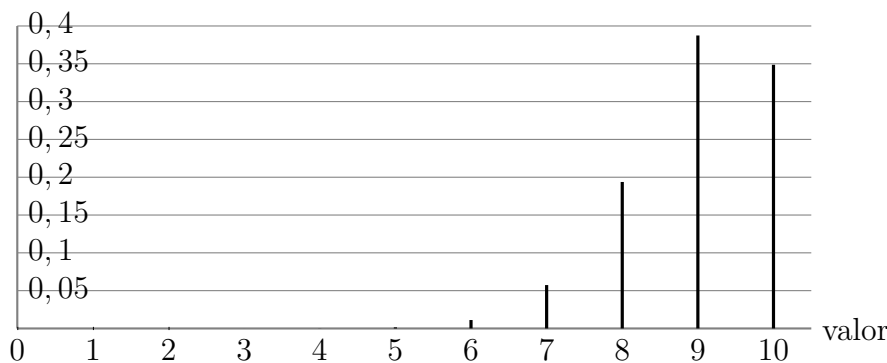
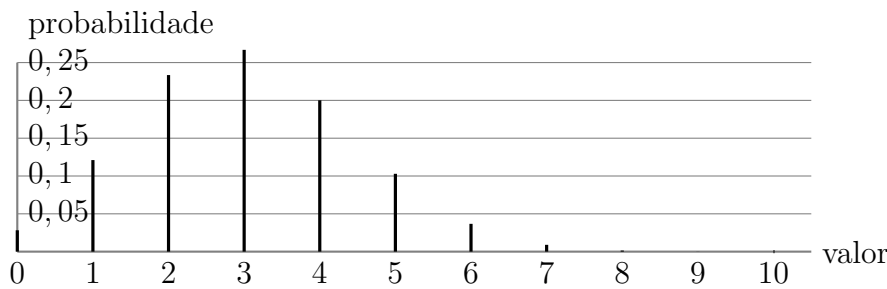
Valores da distribuição binomial com $n=10$ e $p=0,5$:

0.0009765625	0.0097656250	0.0439453125	0.1171875000
0.2050781250	0.2460937500	0.2050781250	0.1171875000
0.0439453125	0.0097656250	0.0009765625	.

Continuamos na **terceira etapa**. Agora devemos assumir que A (a hipótese alternativa) é verdade e achar a distribuição da quantia observada. Neste caso, a distribuição do número de “caras” a serem vistas em 10 lançamentos também é Binomial, mas o valor preciso de p desta distribuição é desconhecido (pois a hipótese não especifica o valor; aliás, é por isto que essa chama-se imprecisa).

Sobre a distribuição da qual estamos interessados, podemos dizer só coisas genéricas, do tipo:

- se p for 0,3, a distribuição seria assim (veja as transparências a seguir);
- se p for 0,9, a distribuição seria assim (veja as transparências a seguir);
- etc.



Na **quarta etapa** da solução, devemos construir a região crítica. Especificamente falando, devemos dividir o conjunto

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$$

em dois subconjuntos: um, chamado de região crítica, e outro que é o complementar desta região (observe: este é o conjunto de todos os valores possíveis da quantia observada).

O sentido de região crítica é: se o valor da quantia observada cair nesta região, rejeitaremos H , quer dizer, respondemos: o resultado mostrou, de acordo com a regra de decisão por nos construída, que a hipótese nula não é verdade, ou seja, que p não é 0,5.

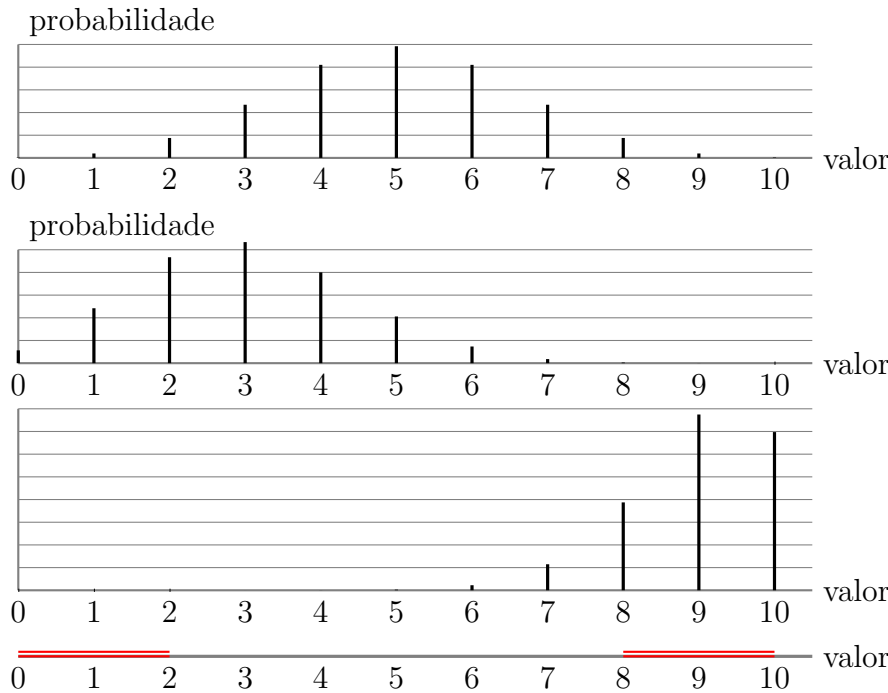
Por outro lado, se o valor da quantia observada cair fora da região crítica, a gente aceitará H , quer dizer, a gente responderá: o resultado mostrou, de acordo com a regra de decisão por nos construída, que a hipótese nula é verdade, quer dizer, que $p = 0,5$.

A divisão do conjunto $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ dá-se pelo princípio de máxima verossimilhança: “ao observar um acontecimento que pode ser causado por duas fontes, acredito naquela das duas que gera este acontecimento com a maior probabilidade”.

Olhando as três distribuições, decide-se que a região crítica deve ser nos extremos do conjunto. Uma das sugestões possíveis é:

$$R_C = \{0, 1, 2\} \cup \{8, 9, 10\}$$

Veja o desenho seguinte para a motivação. Nela, a região crítica está em vermelho.



Uma outra sugestão seria:

$$\{0, 1\} \cup \{8, 9\}$$

Qual das duas é melhor? E é fácil ver que há outras. Qual será a melhor de todas as alternativas possíveis?

No momento, não há critério para a seleção. Mas quando você aprender o significado do nível de significância (ensinado na aula) e da função do poder de teste (não ensinado), você poderá usar estes como critério da seleção. Nem sempre há uma melhor escolha. Podem ser várias. As vezes, a pessoa, que deseja o teste, tem critérios para a escolha do nível de significância e do poder de teste (por exemplo, os limites para o risco de investimento). Neste caso, os valores solicitados determinam a região crítica.

Na **quinta etapa** devemos calcular as probabilidades dos erros. Novamente, o fato da hipótese alternativa ser imprecisa faz com que a probabilidade do erro tipo II seja incalculável. Então, vamos calcular só a do erro tipo I. Recordo: é a probabilidade da regra de decisão de errar quando a verdade está com a hipótese nula, quer dizer, a probabilidade da regra de decisão rejeitar a hipótese nula enquanto na realidade, ela é verdadeira. Isto possui a seguinte expressão matemática:

$$P[X \in R_C] \text{ onde } X \sim Bin(10; 0,5)$$

ou

$$P[X = 0 \text{ ou } 1 \text{ ou } 2 \text{ ou } 8 \text{ ou } 9 \text{ ou } 10]$$

onde $X \sim Bin(10; 0, 5)$.

Eis a tabela da distribuição da variável aleatória binomial com $n = 10$ e $p = 0, 5$.

k	$IP[X = k]$	k	$IP[X = k]$
0	0.0009765625	10	0.0009765625
1	0.0097656250	9	0.0097656250
2	0.0439453125	8	0.0439453125
3	0.1171875000	7	0.1171875000
4	0.2050781250	6	0.2050781250
5	0.2460937500		

De onde conclui-se que

$$IP[X \in R_C] = 0.109375$$

isto é, α , o nível de significância do teste é 0,109375 (aproximadamente 11%).

Neste, o trabalho da construção acaba.

Suponha que fomos fazer uma amostra e observamos 7 caras em 10 lançamentos. Então, diz-se que o teste de hipóteses, cujo nível de significância era 11%, aceitou a hipótese que a moeda é honesta. Ou: o resultado 7 caras em 10 lançamentos aceitou que a moeda é honesta ao nível de significância de 11%.

Este tal nível de significância é a indicação que poderíamos ter errado; para os que entenderam as etapas de construção acima explicadas, o valor 11% significa que o possível erro tinha a probabilidade de acontecer em 11% das vezes.

9.5 O caso quando nenhuma hipótese é precisa

Quando a situação é tal que nenhuma hipótese atribui valor exato ao parâmetro desconhecido, então não se pode calcular nem o nível de significância do teste de hipóteses, nem o seu poder. Isto não significa que o teste não pode ser proferido. Isto simplesmente acarreta que a análise da conclusão fique mais complexa. Tal complexidade não está bem-vinda pelo critério do nível do presente texto. Por esta razão, as situações com duas hipóteses imprecisas são transformadas em situações nas quais uma delas torna-se hipótese precisa. Esta transformação não será por conta de aluno/leitor, pois sua qualificação não permitiria a identificação daquela das hipóteses que deve ser transformada para tornar-se hipótese precisa. Portanto, a transformação, quando necessária, será sempre indicada no enunciado dos exercícios e exemplos do presente texto. Também, muitos exercícios e exemplos do presente texto que refletem situações reais serão artificialmente alterados da maneira que garanta que uma de suas hipóteses é precisa.

- √ **Exemplo 64.** Uma agência governamental está encarregada de fiscalizar a contaminação de um certo produto alimentício, através da análise de uma amostra dos pacotes desse produto. Uma porcentagem de contaminação de até 7% é considerada tolerável. Se a porcentagem de contaminação for maior que este valor o produtor deverá ser autuado. Uma norma dessa agência estabelece que, se no exame de 100 pacotes desse produto forem detectados pelo menos 12 pacotes contaminados, então a fábrica deve ser multada.
- Defina as hipóteses estatísticas adequadas ao problema.
 - Qual é a região crítica escolhida pela agência?
 - Quais são os significados dos erros tipo I e tipo II para o problema?
 - Qual é o nível de significância correspondente à região crítica escolhida?
 - Qual será a decisão da agência se forem observados 10 pacotes contaminados? Dê a resposta completa, quer dizer, as palavras “ao nível de significância” devem estar presentes na resposta; dê a interpretação a esta palavras.

Talvez, seja útil a seguinte visão simplificada da historinha contada acima: Há uma fábrica que produz um certo suplemento em pacotes. Toda a produção está estocada num galpão. Alguns pacotes contêm suplemento contaminado, outros não. O proprietário da fábrica acredita que a proporção de pacotes contaminados (a ser denotada por p) é de no máximo 7% de toda a população de pacotes que ficam dentro de galpão. Uma comissão (enviada por uma agência de controle) acredita que p é maior que 7%. A comissão executará a Regra de Decisão descrita no enunciado. As perguntas (b)-(e) indagam sobre tal regra. A pergunta (a) está aqui para lhe mostrar a transformação de uma das hipóteses que a torna precisa.

É óbvio que as alternativas conflitantes aqui são:

1-a hipótese: $p \leq 7\%$; e

2-a hipótese: $p > 7\%$.

A 1-a é o fabricante está alegando; naturalmente, ele está ciente das normas da agência e por isso controlava sua produção para garantir que a contaminação não ultrapasse 7%. Ao despeito de todo esforço do fabricante, a produção poderia ter saído de seu controle o que, se acontecesse, faria p ser maior que 7%. A comissão controladora veio para verificar qual dos dois cenários ocorreu. É na boca dos membros da comissão que nós colocamos a afirmação $p > 7\%$.

Se formos seguir as etapas de teste de hipóteses descritas na Seção 9.3, perceberemos a impossibilidade de execução da terceira e da quinta etapa, pois em ambas as etapas precisa calcular a distribuição probabilística da quantia observada em duas realidades virtuais: uma é quando uma das hipóteses é válida, e outra – quando a outra. para que poderemos ver isto, precisa-se inicialmente identificar a quantia que está observada na presente situação. Esta é a quantidade de pacotes contaminadas numa amostra de 100 pacotes retiradas aleatoriamente do

galpão (da população de todos os pacotes, quer dizer). Agora podemos ver que

$$\text{a quantia observada tem a distribuição Bin}(100; p) \quad (9.12)$$

Mas como nenhuma das duas hipóteses fala nada específico sobre o valor de p , então não podemos extrair muita informação desta distribuição, e em particular, não podemos calcular nem o nível de significância do teste, nem seu poder.

Existem diversos caminhos que permitem contornar o problema indicado no parágrafo acima. Um caminho bem simples é substituir a hipótese $p \leq 7\%$ pela hipótese $p = 7\%$. Então, a situação com as hipóteses agora é a seguinte:

1-a hipótese: $p = 7\%$; e

2-a hipótese: $p > 7\%$.

No próximo passo de solução, vamos distribuir os nomes, lembrando da seguinte regra: a que é exata será chamada de nula:

hipótese nula (H): $p = 7\%$; e

hipótese alternativa (A): $p > 7\%$.

Com isto, respondemos ao item (a) da tarefa.

Para responder à pergunta do item (b), é preciso lembrar que a Região Crítica é o nome do conjunto de valores tais que se a quantia em observação assumir valor dentro do conjunto, então será aceita a hipótese alternativa. Com isso em mente, olhando o enunciado, deduz-se que Região Crítica é

$$\{12, 13, 14, \dots, 100\} \quad (9.13)$$

De fato, já que o enunciado alega que a hipótese alternativa será aceita caso o número de pacotes contaminados na amostra for pelo menos 12, então a Região Crítica é de 12 adiante. Mas como 100 é o maior valor de pacotes contaminados numa amostra de 100 pacotes, então a Região Crítica acaba em 100.

Observação 1: Apesar de ter já respondido à pergunta do item (b), gostaria de comentar sobre a razão que fez a Agência Governamental estabelecer a Regra de Decisão para a qual achamos agora a sua Região Crítica. Para entender tal razão, é preciso voltar à afirmação (9.12). Ao juntar esta afirmação com o fato de que o aumento do parâmetro p da distribuição binomial acarreta que a variável aleatória com esta distribuição assume maiores valores com probabilidade maiores, concluímos – vai a aplicação do Princípio de Máxima Verossimilhança – que grandes valores da quantia observada deve sugerir a aceitação da hipótese alternativa, enquanto que os seus valores pequenos devem indicar a validade e, portanto, a aceitação da hipótese nula. É por isso que a Região Crítica deve ser composto de valores grandes que a quantia em observação pode assumir.

Observação 2: É óbvio que 7% de 100 dá 7, o que sugere que é 7 o número de pacotes contaminados a serem encontrados em 100 pacotes testados na condição quando a produção enquadra-se nos limites toleráveis. Tal sugestão é aceitável pela nossa intuição, mas não pela lógica da Teoria de Probabilidades. Isto por que apesar de 7 corresponder à média de pacotes contaminados em 100 testados, o número de tais pacotes numa amostra não tem obrigação a coincidir com a média. Então o norma seguida pela agência elevou 7 a 11 com o intuito de não punir a produção correta só por causa do desvio do número amostral da média. Esta elevação é tal de “tolerância da regra de decisão às flutuações aleatórias”.

Prosseguiremos para o item (c). O significado dos erros de tipo I e II foram explicados na Seção 9.3; eu não vou repeti-los agora. Vamos ao cálculo da probabilidade de ocorrer o Erro do tipo I (chamado de nível de significância e denotado tipicamente por α). Assumindo

0	7.051717e-04	20	1.287502e-05
1	5.307744e-03	21	3.691763e-06
2	1.977563e-02	22	9.978225e-07
3	4.862394e-02	23	2.547036e-07
4	8.875176e-02	24	6.150773e-08
5	1.282606e-01	25	1.407403e-08
6	1.528554e-01	26	3.055775e-09
7	1.544990e-01	27	6.303829e-10
8	1.351866e-01	28	1.237042e-10
9	1.040146e-01	29	2.311713e-11
10	7.124438e-02	30	4.117998e-12
11	4.387484e-02	31	6.999026e-13
12	2.449285e-02	32	1.135931e-13
13	1.247940e-02	33	1.761821e-14
14	5.837140e-03	34	2.613201e-15
15	2.518966e-03	35	3.709060e-16
16	1.007248e-03	36	5.040687e-17
17	3.746122e-04	37	6.562708e-18	...	98 2.826045e-110
18	1.300177e-04	38	8.189458e-19	...	99 4.297233e-113
19	4.223552e-05	39	9.799351e-20	...	100 3.234477e-116

a validade da hipótese nula, a quantidade de pacotes contaminados na amostra de 100 pacotes tem distribuição $Bin(100; 0,07)$. Recordo que o nível de significância do teste é a probabilidade da regra de decisão apontar a hipótese nula como falsa enquanto que esta é verdadeira. Em símbolos matemáticos:

$$\alpha = \mathbb{P}[X \geq 12] \text{ onde } X \sim Bin(100; 0,07)$$

Usando a tabela dos valores desta distribuição binomial (veja a tabela), conclui-se que

$$\alpha = 1 - \mathbb{P}[X \leq 11] = 1 - 0.9530996 \approx 0,05(5\%)$$

O cálculo do nível de significância pode ser feito com auxílio da aproximação de binomial pela normal, o que é admissível no caso já que a distribuição de binomial tem n grande (100) e tem np maior que 4. (Na verdade, $np = 100 \times 0,07 = 7$ o que não é muito grande, e, conseqüentemente, não se espera conseguir uma aproximação muito exata.) Para a aproximação, precisamos da esperança e da variância da variável aproximada; eis estas:

$$X \sim Bin(100; 0,07) \implies \mathbb{E}[X] = 100 \times 0,07, \quad \text{Var}[X] = 100 \times 0,07 \times 0,93$$

conseqüentemente,

$$\alpha = \mathbb{P}[X \geq 12] \approx \mathbb{P}[Y \geq 12], \quad \text{onde } Y \sim \mathcal{N}(7; 6,51)$$

Fazendo a padronização e usando a Tabela da Distribuição Normal Padrão (na penúltima passagem dos cálculos abaixo apresentados), temos (abaixo, Z denota a variável aleatória Normal Padrão):

$$\alpha \approx \mathbb{P}[Y \geq 12] = \mathbb{P}\left[Z \geq \frac{12 - 7}{\sqrt{6,51}}\right] \approx \mathbb{P}[Z \geq 1,96] = 1 - 0,9750 = 0,025$$

Agora é a vez da resolução do item com a última pergunta: “Qual será a decisão da agência se forem observados 10 pacotes contaminados?” Já que 10 está fora da Região Crítica, então deve ser aceita a hipótese nula. Este é o cerne da resposta. Entretanto, para fins didáticos, devemos colocar a resposta em sua forma completa e detalhada. Eis esta: “Com o resultado “10 pacotes contaminados na amostra de 100 pacotes retiradas de população”, o teste ha hipótese nula “a proporção de pacotes contaminados (p) na população é 0,07” contra a hipótese alternativa “ $p > 0,07$ ” com a nível de significância de 0,05 aceitou a hipótese nula.”Fim do **Exemplo 64**↑

9.6 O tratamento genérico para duas hipóteses imprecisas

9.7 Nível discretivo ou P-valor

Nesta seção estão definidos os conceitos “nível discretivo” e “P-valor” para o Teste de Hipóteses sobre Proporção. A apresentação está feita da forma que permite ao leitor estender as definições para qualquer outro tipo de teste de hipóteses. Vale avisar que os conceitos “nível discretivo” e “P-valor” estão muito usados na prática de Testes de Hipóteses. Inclusive, muitos pacotes estatísticos emitem suas respostas nestes termos.

Imagine uma urna com bolas pretas e brancas. Denote por p a proporção de bolas pretas na urna. Suponha que seu valor numérico é desconhecido e que a respeito dele há dúvida que está formalizada em (9.14) abaixo no formato de concorrência de duas hipóteses:

$$\begin{aligned} \text{hipótese nula } (H) & : p = 0,4 \\ \text{e hipótese alternativa } (A) & : p > 0,4 \end{aligned} \quad (9.14)$$

Suponha que está decidido que a escolha entre as duas hipóteses será feita com base no número de bolas pretas a serem vistas na amostra aleatória de $n = 100$ bolas a serem retiradas da urna. Denotamos este número por X . Na perspectiva de que a amostragem acontecerá no futuro, tal quantia é vista como variável aleatória, fato que justificou a notação da quantia por letra maiúscula (recorde que no presente texto, letras maiúsculas denotam variáveis aleatórias). Recorde que os argumentos das seções anteriores permitem facilmente deduzir a distribuição de X ; ela é $\text{Bin}(100; p)$. No que se segue, vamos usar não esta distribuição mas a distribuição que a aproxima; ela é $\mathcal{N}(100p; 100p(1-p))$ (aqui usamos o fato de que a distribuição binomial pode ser bem aproximada pela Normal desde que sua média e sua variância coincidam com, respectivamente, a média e a variância da distribuição binomial). Concluindo, temos que

$$\text{a quantidade de bolas pretas na amostra segue } X \sim \mathcal{N}(100p; 100p(1-p)) \quad (9.15)$$

Imagine ainda que será aplicada a seguinte

Regra de Decisão: caso (após tomar a amostra) X apresentar valor menor que certo limiar ℓ , decidir-se-á que H é verdadeira, e se o valor for ℓ ou maior, decidir-se-á que A é verdade.

Com o intuito de mostrar o sentido do conceito *nível discretivo*, eu preciso primeiramente calcular regiões críticas para alguns valores do nível de significância (α). Para tal, recorde da Teoria Geral de Teste de Hipóteses que, a Região Crítica no presente caso é $[\ell, \infty)$ e que

$$\alpha = \mathbb{P}[X \geq \ell], \text{ onde } X \sim \mathcal{N}(100 \times 0,4; 100 \times 0,4(1 - 0,4)) \quad (9.16)$$

(recorde que o cálculo do nível de significância se faz sob a “realidade virtual” de que H é verdadeira, fato que fez a substituição de p em (9.15) por $0,4$). O caminho do cálculo passa pela normalização da variável aleatória X :

$$\alpha = \mathbb{P}\left[Z \geq \frac{\ell - 40}{\sqrt{24}}\right], \text{ onde } Z \text{ é Normal Padrão} \quad (9.17)$$

A fórmula (9.17) e a Tabela da Distribuição Acumulada de Normal Padrão dão os resultados apresentados na tabela da Figura 9.7, em suas duas primeiras colunas.

Suponha que amostragem foi feita e foram observadas 50 bolas pretas na amostra. Isto quer dizer que X assumiu valor 50. A notação genérica para o valor assumido e para X é x_{obs} . No presente caso, então, $x_{obs} = 50$. Nossa corrente tarefa é analisar a conclusão de Regra de Decisão em posse deste x_{obs} . Naturalmente, tal conclusão depende da Região Crítica a qual, por sua vez, depende do nível de significância com o qual desejamos deduzir a conclusão. Tudo

Nível de significância	Valor do limiar da Região Crítica	A conclusão da Regra de Decisão com a observação $x_{obs} = 50$
0,115	46	$48 < 52 \Rightarrow$ aceita-se H
0,087	47	$47 < 50 \Rightarrow$ aceita-se H
0,055	48	$48 < 50 \Rightarrow$ aceita-se H
0,036	49	$49 < 50 \Rightarrow$ aceita-se H
0,023	50	$50 = 50 \Rightarrow$ rejeita-se H
0,014	51	$51 > 50 \Rightarrow$ rejeita-se H
0,008	52	$52 > 50 \Rightarrow$ rejeita-se H
0,005	53	$53 > 50 \Rightarrow$ rejeita-se H

Figura 9.7: Alguns pares de nível de significância–limiar da Região Crítica e as correspondentes decisões da Regra de Decisão para o resultado x_{obs} .

isto está apresentado na tabela da Figura 9.7. A última coluna da tabela é a solução da tarefa agora colocada.

Ao analisar a última coluna da tabela, percebe-se que quando o limiar é menor que x_{obs} então a Regra decide a favor do hipótese nula, mas quando o limiar ultrapassa x_{obs} aí então a Regra começa a decidir a favor da hipótese alternativa e continua com tla “preferencia” para todos os limiares maiores de x_{obs} . Esta dinâmica da decisão da Regra pode ser formulada em termos de nível de significância. Para tal, vamos dar o nome **P-valor**, ou, **nível descritivo** ao nível de significância correspondente ao valor x_{obs} do limiar. Com isto podemos descrever da seguinte maneira

A dinâmica da decisão da Regra: para qualquer valor do nível de significância maior de que o P-valor a Regra de Decisão vai aceitar a hipótese nula, já quando o valor do nível de significância for igual a ou maior de P-valor, então a Regra vai rejeitar a hipótese nula.

A conclusão do parágrafo acima revela para nos a utilidade do conceito P-valor. Eis esta: no diálogo entre estatístico que executa teste de hipóteses e o cientista que trouxe os dados e está interessado nos resultados do teste, o estatístico pode dizer assim: “O P-valor é tal, e com este você sabe o resultado de teste para cada valor de nível de significância que você escolher.” Concorde que esta resposta é mais eficiente do que a forma de interação entre estatístico e cientista segundo a qual o cientista deve trazer para o estatístico o nível de significância desejado para que esse consiga emitir a decisão de teste.

As seguintes observações não são secundárias aos conceitos discutidos na presente seção.

Observação 1. É importante notar que P-valor está associado ao valor observado da quantia usada na Regra de Decisão. Por isto, nas aulas de Estatística Básica eu falo “nível descritivo associado/correspondente ao valor observado” e “P-valor do valor observado” no lugar de tradicionais termos “nível descritivo” e “P-valor”.

Observação 2. Segue das definições da presente seção que no caso considerado

$$\text{P-valor da observação } x_{obs} \text{ é } \mathbb{P}[X \geq x_{obs}] \text{ com a distribuição de } X \text{ determinada pela } H \quad (9.18)$$

Esta fórmula confunde pessoas, pois por um lado ela insinua que a quantidade de bolas pretas na amostra deve ser vista como variável aleatória X , o que implica no que todo o cálculo da probabilidade é feito antes de colher amostra, mas, por outro lado, a presença de x_{obs} na fórmula indica que o cálculo só pode ser feito após ter colhido amostra. A interpretação correta mistura

as duas visões. A interpretação é: “A probabilidade do número de bolas pretas ser igual a ou maior que o valor observado na amostra”.

Observação 3. Se você, meu leitor precisa interpretar P-valor ou usá-lo para escolher entre hipóteses, então suas necessidades são plenamente atendidas pela presente apresentação. Já se você pretende calcular P-valores, saiba que a técnica de cálculo está apresentada, mas que há detalhes que não foram trabalhados mas aos quais você precisa prestar atenção quando fazer seus cálculos. Tais detalhes dependem do tipo de teste, da formulação de hipóteses e da Regra de Decisão. Por exemplo, se no Teste de Hipóteses para Proporção aqui considerado a hipótese alternativa alegasse $p < 0,4$ em vez de $p > 0,4$, então a desigualdade “ \geq ” nas relações (9.16), (9.17) e (9.18) deveria ser trocada por “ \leq ”.

Observação 4. A relação (9.18) permite dizer que P-valor é a probabilidade da amostra mostrar o valor observado ou qualquer outro valor menos típico que o observado na perspectiva da validade da hipótese nula. Isto sugere interpretar P-valor como *a força com qual o valor observado indica à validade da hipótese nula*. Consequentemente, pode pensar assim: se P-valor é pequeno tal força é fraca e, consequentemente, a hipótese nula deve ser rejeitada, já se P-valor for grande então a hipótese nula deve ser aceita. É importante notar que “grande” e “pequeno” não tem graduação absoluta. O custo e o perigo associado à aceitação errada da hipótese nula é um dos fatores principais que determina a magnitude de “pequeno” e “grande”. Por isto que ela será bem diferente em, por exemplo, pesquisa de opinião e em estudos clínicos de remédio.

9.8 Um exemplo para auto-teste de sua compreensão de princípios de teste de hipóteses

√ **Exemplo 65.** Sabe-se que há dez anos a população paulistana era constituída de 15% de gaúchos. Um pesquisador alega que essa população aumentou, pois a migração gaúcha é uma das maiores do país. Para testar sua teoria, ele pretende coletar uma amostra de 60 pessoas residentes na cidade e calcular a proporção dos gaúchos nesta. A escolha entre as hipóteses H : a porcentagem p de gaúchos na população paulistana é exatamente 15%;
 A : a porcentagem p de gaúchos na população paulistana é maior que 15%;
 será feita com base no resultado a ser obtido por pesquisador.

(Qa) Queremos que o nível de significância do teste seja 7%. Qual é a Região Crítica do teste (use a distribuição binomial em anexo)?

(Qb) Qual será a resposta do teste, se 11 das 60 foram gaúchos?

(Qc) Refaça as contas da Região Crítica caso a amostra for de 600 pessoas (queremos que o nível de significância continue ser 7%); agora use a aproximação de binomial pela normal. Qual será a conclusão do teste caso 110 dos 600 forem gaúchos?

Pergunta 1. As hipóteses versam sobre a proporção de que? Ou, em outras palavras, o símbolo p que figura nas hipóteses, refere-se ao que? Observe, que ao total, são três possibilidades:

- (1) a proporção dos gaúchos na população paulistana há dez anos;
- (2) a proporção dos gaúchos na população paulistana hoje;
- (3) a proporção dos gaúchos na amostra.

Uma e só uma dessas três possibilidades é a resposta correta na pergunta. Qual é?

Pergunta 2. Por que H afirma um valor exato sobre p ? Como o enunciado foi montado para permitir isso? Se tirasse do enunciado a parte que permite dar um valor exato para p na hipótese H , como seria a formulação dessa? Tal mudança do enunciado afetaria a formulação do hipótese A ? Se sim, então como? Senão, então por que não?

Pergunta 3. Voltamos para a condição do enunciado e para as formulações das hipóteses de acordo com o enunciado. Para essa forma das hipóteses, qual deve ser o formato da Região Crítica do teste que vai poder “decidir” entre essas com base no número de gaúchos na amostra de 60 paulistanos? Experimente diversas formas para a Região Crítica e revele as suas vantagens e desvantagens, sendo julgadas pela capacidade e eficiência na “decisão” entre as hipóteses.

Pergunta 4. Tome a Região Crítica no seu formato correto (o qual você deduziu ao responder a Pergunta 3). Mude um dos extremos da Região (você vai ver facilmente, que há um só extremo da Região cuja mudança faz sentido). Como tal mudança afeta o nível de significância do teste? Fix a Região Crítica que corresponde ao nível de significância de 7% e responda na pergunta do enunciado.

Solução para questões (Qa)-(Qc) de Exemplo 65. Seja p a proporção de gaúchos na população paulistana e X a v.a. *número de gaúchos em São Paulo*. O teste será realizado para:

- $\mathbb{H}_0 : p = 0,15$;
- $\mathbb{H}_A : p > 0,15$

A região crítica será determinada por um valor c tal que:

$$\mathcal{RC} = \{X : X \geq c\}$$

k	$\mathbb{P}[X = k]$	$\mathbb{P}[X \leq k]$	$\mathbb{P}[X \geq k - 1]$
0	5.822849e-05	5.822849e-05	1.000000e+00
1	6.165370e-04	6.747655e-04	9.999418e-01
2	3.209619e-03	3.884384e-03	9.993252e-01
3	1.095046e-02	1.483485e-02	9.961156e-01
4	2.753720e-02	4.237205e-02	9.851652e-01
5	5.442646e-02	9.679851e-02	9.576280e-01
6	8.804281e-02	1.848413e-01	9.032015e-01
7	1.198566e-01	3.046979e-01	8.151587e-01
8	1.401265e-01	4.448244e-01	6.953021e-01
9	1.428740e-01	5.876984e-01	5.551756e-01
10	1.285866e-01	7.162850e-01	4.123016e-01
11	.031444e-01	8.194294e-01	2.837150e-01
12	7.432461e-02	8.937540e-01	1.805706e-01
13	4.842871e-02	9.421827e-01	1.062460e-01
14	2.869096e-02	9.708737e-01	5.781730e-02
15	1.552687e-02	9.864005e-01	2.912634e-02
16	7.706351e-03	9.941069e-01	1.359947e-02
17	3.519856e-03	9.976267e-01	5.893119e-03
18	1.483861e-03	9.991106e-01	2.373263e-03
19	5.788435e-04	9.996894e-01	8.894024e-04
20	2.094052e-04	9.998988e-01	3.105588e-04
21	7.038829e-05	9.999692e-01	1.011537e-04
22	2.201987e-05	9.999913e-01	3.076538e-05
23	6.420114e-06	9.999977e-01	8.745519e-06
24	1.746649e-06	9.999994e-01	2.325405e-06
25	4.438543e-07	9.999999e-01	5.787564e-07
26	1.054405e-07	1.000000e+00	1.349021e-07
27	.343122e-08	1.000000e+00	2.946163e-08
28	.873300e-09	1.000000e+00	6.030405e-09
29	9.489590e-10	1.000000e+00	1.157105e-09
30	1.730455e-10	1.000000e+00	2.081456e-10
31	2.955236e-11	1.000000e+00	3.510025e-11
32	4.726204e-12	1.000000e+00	5.547880e-12
33	7.076670e-13	1.000000e+00	8.216869e-13
34	9.917132e-14	1.000000e+00	1.139373e-13
35	1.300061e-14	1.000000e+00	1.477697e-14
36	1.593212e-15	1.000000e+00	1.815257e-15
37	1.823709e-16	1.000000e+00	1.823709e-16
38	.947924e-17	1.000000e+00	1.947924e-17
39	1.939110e-18	1.000000e+00	1.939110e-18
40	1.796528e-19	1.000000e+00	1.796528e-19

Figura 9.8: Valores relacionados a $X \sim \text{Bin}(60; 0, 15)$. Na coluna à esquerda, são valores de 0 a 40 que X pode assumir; na realidade X assume valores inteiros entre 0 e 60, mas devido ao limite de espaço, apresentamos somente os de 0 até 40. Na segunda coluna, apresentamos $\mathbb{P}[X = k]$, na terceira – $\mathbb{P}[X \leq k]$, e na última – $\mathbb{P}[X > k]$ o que é a mesma coisa que $\mathbb{P}[X \geq k - 1]$.

Fixando $\alpha = 7\%$, podemos determinar o valor de c da seguinte forma:

$$\begin{aligned}
 \alpha &= \mathbb{P}(\text{erro I}) \\
 0,07 &= \mathbb{P}(\text{rejeitar } \mathbb{H}_0 | \mathbb{H}_0 \text{ é verdadeira}) \\
 0,07 &= \mathbb{P}(X \in \mathcal{RC} | p = 0, 15) \\
 0,07 &= \mathbb{P}(X \geq c | p = 0, 15)
 \end{aligned}
 \tag{9.19}$$

Sob a hipótese nula, $X \sim \text{Binomial}(60; 0,15)$. Procuramos então na tabela desta distribuição o valor c tal que $\mathbb{P}[X \geq c] = 0,07$. A tabela (na sua terceira coluna) não contém valor 0,07. O valor mais próximo é 0,05781 que corresponde a $\mathbb{P}[X \geq 14 - 1]$, isto é, o procurado c é 13, e, portanto, a Região Crítica correspondente é $\{13, 14, \dots, 60\}$. Com a observação 11 devemos então aceitar a hipótese nula, já que 11 não está na Região Crítica. Observe que a resposta completa é: “Ao nível de significância 7% aceitamos a hipótese nula”.

Se a amostra for de 600 e não de 60 pessoas, então $X \sim \text{Binomial}(600; 0,15)$. Podemos aproximar X pela v.a $Y \sim \mathbb{N}(np; np(1-p))$, ou seja, $Y \sim \mathbb{N}(90; 76,5)$. Dessa forma, (9.19) pode ser aproximada por:

$$\begin{aligned} 0,07 &= \mathbb{P}(Y > c) \\ 0,07 &= \mathbb{P}\left(\frac{Y - np}{\sqrt{np(1-p)}} > \frac{c - 90}{\sqrt{76,5}}\right) \\ 0,07 &= \mathbb{P}\left(Z > \frac{c - 90}{\sqrt{76,5}}\right) \end{aligned}$$

Consultando a tabela da distribuição normal padrão, teremos:

$$\begin{aligned} \frac{c - 90}{\sqrt{76,5}} &\approx 1,48 \\ c &\approx 1,48(\sqrt{76,5}) + 90 \\ c &\approx 102,95 \end{aligned}$$

Portanto

$$\mathcal{RC} = [102,95; 600]$$

Se a amostra coletada apresentar um valor $x_{obs} = 110$, temos então que $x \in \mathcal{RC}$ e portanto, para um nível de significância $\alpha = 7\%$, rejeitamos \mathbb{H}_0 , ou seja, há evidências de que a proporção de gaúchos na população paulistana aumentou.

Fim do Exemplo 65↑

9.9 Teste de hipóteses versus estimação intervalar

9.9.1 Teste de proporção versus estimação intervalar

Neste exemplo, nos vamos analisar a afirmação de um político de que

$$\text{seu eleitorado é no mínimo 60\% da população com a confiança de 90\%} \quad (9.20)$$

e a análise basear-se-á no seguinte dado:

das 100 pessoas escolhidas ao acaso da região do eleitorado, 58 afirmaram seu apoio ao político. (9.21)

Antes de começar, gostaria de traduzir este exemplo para a linguagem usada no Exemplo 1.

Eis a tradução. Vamos representar cada pessoa do eleitorado por uma bolinha. Todas as bolinhas têm o mesmo tamanho, mas as que correspondem aos apoiadores do político estão pintadas de preto; outras são brancas. Escolher 100 pessoas ao acaso do eleitorado e contar o número destas que apoiam o político é a ação equivalente ao seguinte procedimento: colocar todas as bolas numa urna, misturá-las e, ao sortear desta 100 bolas com reposição, contar o número das pretas dentro das sorteadas. Cabe um “porém” à equivalência alegada: a seleção de pessoas não ocorre de pleno acordo com o sorteio de bolas pois as pessoas já selecionadas não serão selecionadas novamente, enquanto que no sorteio com reposição, uma bola pode ser retirada da urna por mais que uma vez. Entretanto, o desacordo será desconsiderado e a razão para tal é a insignificância dele quando a população do eleitorado é muito grande em comparação com o tamanho da amostra a ser feita (que é 100). Neste caso, podemos assumir a retirada de até 100 pessoas da população – quaisquer que seja o número dos apoiadores do político dentro das retiradas – não altera a proporção dos apoiadores na população restante. Consequentemente, ao denotar esta proporção por p , podemos alegar que – sob este pressuposto que assumimos incondicionalmente e que não serão mais lembrado em cada linha de texto – ocorre que

- (a) em qualquer uma das 100 escolhas, a probabilidade que a pessoa escolhida seja o eleitor do político é p ,
- (b) e a probabilidade referida no item acima não depende dos resultados das escolhas de outras 99 pessoas (aqui, o resultado é o rótulo para “a pessoa é eleitor do político” e para “a pessoa não é eleitor do político”).

Agora, voltando à nossa urna com bolas, podemos alegar que o sorteio de 100 bolas da urna com a reposição apresenta as seguintes características:

- (a') a probabilidade de qualquer uma das 100 bolas sorteadas ser de cor preta é p
- (b') e a probabilidade referida no item acima não depende dos resultados observados para bolas sorteadas em outros sorteios.

Há uma clara equivalência entre (a),(b) e (a'), (b'). Ela nos permite reformular a situação do Exemplo 2 em termos de urna, bolas e sorteio e discutir nestes termos as questões surgidas nesta e suas soluções. As vantagens desta reformulação serão discutidas mais para o final do texto todo.

Então, o político afirma que

$$\text{a proporção de bolas pretas na urna é de no mínimo 60\%} \quad (9.22)$$

e o objetivo do corrente exemplo é confirmar ou rejeitar esta (com base nos dados da amostra). Para que este objetivo seja executado de acordo com a abordagem clássica de testes de hipótese, introduzimos a afirmação que nega a do político:

$$\text{a proporção de bolas pretas na urna é menor que } 60\% \quad (9.23)$$

e colocamos o objetivo como

$$\text{escolher, com base nos dados de amostra, entre as hipóteses (9.22) e (9.23)} \quad (9.24)$$

Ainda mais, para que possamos dar uma apresentação menos formal que a tradicional de livros didáticos sobre métodos estatísticos, vamos introduzir na nossa historinha a figura de um jornalista que contesta a afirmação do político, e a figura de um estatístico, que será a pessoa da historinha atarefado a decidir qual dos dois, o político ou o jornalista, tem mais razões. Recordamos, antes de ingressar no caminho que nos levará ao objetivo colocado, que o termo *hipótese* substituiu o termo *afirmação* por causa de a probabilidade de obter um apoiador do político ao entrevistar qualquer uma das 100 pessoas a serem escolhidas com bolas na urna nos queremos modelar as duas principais características da seleção de pessoas, as quais são: , diferença A frase “60% da população votam em mim” da afirmação do político significa que seu eleitorado pode ser representado por uma urna com bolinhas de mesmo tamanho

9.9.2 Teste qui-quadrado versus estimações intervalares

As estimações intervalares parecem ter a capacidade de competir com Teste Qui-quadrado, quando esse aplica-se para verificar a aderência. Vamos ver o exemplo abaixo que nos mostrará como acontece a competição e porquê essa não é justa. O enunciado do exemplo é uma dos exercícios da lista sobre Teste qui-quadrado.

Exemplo. Um modelo de automóvel é vendido em quatro versões: SX, LX, GLX, GTX. Foi feita uma campanha publicitária para melhorar as vendas das versões GLX e GTX. Posteriormente, foi verificada a escolha das versões em 500 vendas escolhidas ao acaso. Os resultados foram:

Versão	SX	LX	GLX	GTX	total
Unidades vendidas	210	125	105	60	500

De acordo com o fabricante, a participação de cada versão nas vendas deste modelo até a realização da campanha era 40% de SX, 30% de LX, 20% de GLX e 10% de GTX.

(a) Utilize um teste de hipóteses adequado para verificar se houve ou não mudanças na participação de cada versão nas vendas após a campanha. Indique o (nome do) teste utilizado, escreva as hipóteses correspondentes e especifique os graus de liberdade. Calcule o nível descritivo (valor $-p$) para os dados observados, e conclua adotando o nível de significância de 5%.

Estamos interessados em testar as hipóteses:

H_0 : A campanha publicitária não teve efeito nas vendas das versões desse modelo de automóvel.

H_1 : A campanha publicitária teve efeito nas vendas das versões desse modelo de automóvel.

Ou equivalentemente,

H_0 : $p_{SX} = 0,40$, $p_{LX} = 0,30$, $p_{GLX} = 0,20$ e $p_{GTX} = 0,10$.

H_1 : Existe pelo menos uma diferença.

em que p_{SX} , p_{LX} , p_{GLX} e p_{GTX} representam a participação nas vendas da versão SX, LX, GLX e GTX, respectivamente.

Para testar essas hipóteses precisamos realizar um **teste de aderência**, ou seja, precisamos calcular o valor da estatística de teste

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \stackrel{SobH_0}{\sim} \chi_{(q=k-1)}^2; E_i = np_i.$$

A tabela seguinte apresenta as frequências observadas (O_i), as proporções esperadas (p_i) e as frequências esperadas (E_i) para as quatro versões de automóveis: Observe que $E_i > 5\forall i =$

Versão	Frequência Observada (O_i)	Proporção esperada (p_i)	Frequência esperada (E_i)
SX	210	$p_1 = 0,4$	$E_1 = n \times p_1 = 500 \times 0,4 = 200$
LX	125	$p_2 = 0,3$	$E_2 = n \times p_2 = 500 \times 0,3 = 150$
GLX	105	$p_3 = 0,2$	$E_3 = n \times p_3 = 500 \times 0,2 = 100$
GTX	60	$p_4 = 0,1$	$E_4 = n \times p_4 = 500 \times 0,1 = 50$
Total	500	1	500

1, 2, 3, 4 e n é suficientemente grande. Assim, temos:

$$\begin{aligned} \chi^2 &= \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} = \frac{(210 - 200)^2}{200} + \frac{(125 - 150)^2}{150} + \frac{(105 - 100)^2}{100} + \frac{(60 - 50)^2}{50} \\ &= 0,5000 + 4,1667 + 0,2500 + 2,0000 = 6,9167. \end{aligned}$$

Usando a distribuição Qui-quadrado com $q = 4 - 1 = 3$ graus de liberdade, o nível descritivo é,

$$\text{valor-p} = P(\chi_q^2 \geq \chi_{obs}^2) = P(\chi_3^2 \geq 6,9167) = 0,0746$$

O valor-p pode ser obtido no *Rcmdr*: clique no menu Distribuições → Distribuições Contínuas → Distribuição Qui-Quadrado → Probabilidades da Qui-Quadrado. Em Valores da variável coloque 6,9167, em Graus de liberdade coloque 3 (4-1) e clique em cauda superior.

```
> pchisq(c(6.9167), df=3, lower.tail=FALSE)
[1] 0.07460076
```

Decisão e conclusão: como o nível descritivo do teste é maior do que o nível de significância considerado de 5%, então não rejeitamos H_0 . Ou seja, concluímos que não temos evidências suficientes de que a campanha publicitária teve efeito nas vendas desse modelo de automóvel.

(b) Usando o resultado da pesquisa (o resultado que está apresentado na tabela acima), obtenha uma estimativa intervalar, com o coeficiente de confiança de 95%, para a proporção de compradores, deste modelo, que preferem a versão LX.

Uma estimativa pontual para p_{LX} é $\hat{p}_{LX} = \frac{125}{500} = 0,25$. Temos que $\gamma = 0,95 \rightarrow z = 1,96$. Portanto,

$$IC(p_{LX}; \gamma) = \left[\hat{p}_{LXobs} \pm z \sqrt{\frac{\hat{p}_{LXobs}(1 - \hat{p}_{LXobs})}{n}} \right]$$

$$IC(p_{LX}; 95\%) = \left[0,25 \pm 1,96 \times \sqrt{\frac{0,25(1 - 0,25)}{500}} \right] = [0,25 \pm 1,96 \times 0,019365]$$

$$= [0,25 \pm 0,037955] = [0,2120; 0,2880].$$

Observe que $p_{LX} = 0,25$ encontra-se dentro do intervalo construído, indicando que a participação nas vendas da versão LX realmente não se alterou após a campanha publicitária.

9.10 Teste de Hipóteses para média

Imagine que há uma população de pessoas de alturas diferentes. A presente exposição é sobre a altura média populacional que denota-se por μ e calcula-se como

$$\frac{1}{N} (a_1 + a_2 + \dots + a_N) \quad (9.25)$$

onde N é o tamanho da população e onde a_1, \dots, a_N são as alturas de seus indivíduos. Imagine que μ é desconhecido e que não há condições para medir a altura de cada indivíduo para que a média da altura possa ser calculada via (9.25). Imagine que a única informação que possuímos são as alturas x_1, x_2, \dots, x_n de n indivíduos retiradas da população aleatoriamente e com reposição. Eu pretendo mostrar na presente seção como se usa a amostra x_1, x_2, \dots, x_n para decidir entre as hipóteses

$$\begin{aligned} H : \mu &= \mu_0 \\ A : \mu &> \mu_0 \end{aligned} \quad (9.26)$$

onde μ_0 é certo valor numérico (pode pensar que μ_0 é valor “trazido” pelas pessoas que formularam as hipóteses e solicitaram que seja apontada a mais razoável delas). O procedimento a ser mostrado chama-se *Teste de hipóteses para média de uma população*.

A decisão emprega a méida amostral

$$\frac{1}{n} (x_1 + x_2 + \dots + x_n) \text{ a ser denotada por } \bar{x} \quad (9.27)$$

e a variância amostral

$$\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \text{ a ser denotada por } s^2 \quad (9.28)$$

na seguinte Regra de Decisão: caso

$$\frac{\bar{x} - \mu_0}{\sqrt{s^2/n}} \text{ a ser denotada por } t_{obs} \quad (9.29)$$

estiver na região $[\ell, +\infty)$ (aqui, ℓ é um número real, não necessariamente inteiro), chamada Região Crítica do teste, então aceita-se a hipótese alternativa, caso contrário, aceita-se a hipótese nula.

A Região Crítica está vinculada ao nível de significância do teste via a distribuição chamada t de Student com $n - 1$ graus de liberdade, sob a condição (9.32) a ser formulada em seguida; denotando por T a variável aleatória que tem esta distribuição, tem-se que

$$\text{o nível de significância (denotado por } \alpha) \text{ é } \mathbb{P} [T \geq \ell] \quad (9.30)$$

Os valores da distribuição t de Student estão tabulados; com uso de tabela, se resolvem as duas questões básicas associadas à aplicação do teste: dado o nível de significância desejado achar a correspondente Região Crítica, e dado a Região Crítica usada para decidir acerca da validade de hipótese nula achar o valor do nível de significância desta decisão.

O envolvimento da distribuição t de Student ocorre pela razão que será explicada agora. Para o fim da explicação, introduziremos variáveis aleatórias

$$X_1, X_2, \dots, X_n \quad (9.31)$$

que expressam os valores da amostra na perspectiva de “amostra vai acontecer”. Com a mostra será feita com reposição, então as variáveis aleatórias são independentes. Também, elas todas

têm a mesma distribuição que é a distribuição da altura pela população. Acontece que que não conhecemos esta distribuição, pois se conhecessemos então não haveria pergunta sobre a média. Entretanto, podemos fazer algumas suposições genéricas sobre esta. Se suposemos que

$$\text{a altura tem distribuição normal por indivíduos da população} \quad (9.32)$$

então cada X_i tem a mesma distribuição e conseqüentemente (pelos cálculos complexos que não serão mostrados) tem-se que

$$\frac{\bar{X} - \mu_0}{\sqrt{\frac{1}{n} \left\{ \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \right\}}}, \text{ onde } \bar{X} = \frac{1}{n} (X_1 + \dots + X_n) \quad (9.33)$$

tem a distribuição t de Student com $n - 1$ graus de liberdade. Esta é a razão pela qual o Teste de Hipóteses baseia-se na distribuição t de Student.

9.11 Exercícios do tema Teste de Hipóteses para Proporção

9.11.1 Exercícios sobre os testes nos quais ambas as hipóteses são precisas

Comentários sobre os Exercícios 142–153. *Em relação ao tema “Teste de Hipóteses para Proporção” há três tipos de problemas que posso e vou perguntar de meus alunos/leitores. Um dos problemas é trivial, mas outros dois são complexos no sentido de que suas soluções requerem um bom conhecimento dos princípios do método de Teste de Hipóteses assim como dos cálculos vinculados ao método. Eu achei que a exposição desses problemas e de suas soluções fique satisfatoriamente claro caso tudo for treinado no arcabouço do Exemplo sobre Piscina de Bolinhas. Tal treinamento está oferecido nos Exercícios 142–153.*

Antes de você, meu leitor, partir para a solução desses exercícios, permita-me lhe fornecer a seguinte informação auxiliar acerca dos mesmos. As perguntas dos itens (a) dos Exercícios 142–147 pedem calcular o nível de significância que corresponde à Região Crítica da Regra de Decisão formulada nos seus enunciados. Já os itens (a) dos Exercícios 148–153 trabalham no sentido contrário: eles fornecem o valor de nível de significância e pedem achar a correspondente Região Crítica da Regra de Decisão. Esses dois tipos de tarefa correspondem àqueles dois problemas que foram chamados de complexos no parágrafo acima. Você recorda que além daqueles dois, houve mais um problema chamado de trivial. Esse aparece nos itens (b) de todos os exercícios. A tarefa desses itens é apontar na hipótese a ser aceita pela Regra de Decisão dada a informação sobre a Região Crítica da regra e o resultado de amostra. Tal resultado é nada mais e nada menos de que o número de bolas pretas vistas na amostra (isso é válido para o caso de Piscina de Bolinhas, mas será válido também para qualquer outro problema do tema Teste de Hipóteses para Proporção caso você interpretar o problema em termos de bolas na urna). A resposta é direta: caso o número de bolas pretas na amostra estar dentro da Região Crítica, deve-se então rejeitar a hipótese nula, caso contrário, deve-se aceitar a hipótese nula. Eu repeti aqui essa regra de resposta mirando dois objetivos: primeiramente, justificar o porquê o problema foi chamado de trivial, e, em segundo lugar, recordar para você um dos princípios de nomenclatura ligada ao tema: a Região Crítica é, pela própria definição, o conjunto tal que se o número de bolas pretas de amostra cair nessa região então você deve rejeitar a hipótese nula. Em outras palavras, a Região Crítica está “vinculada” à hipótese nula pela rejeição. Esse lembrete não é excessivo pois para iniciantes torna-se difícil recordar que a Região Crítica “reprova” a hipótese nula e não a hipótese alternativa.

Exc. 142. No Exemplo sobre Piscina de Bolinhas, atribua o nome “hipótese nula” à afirmação do administrador do buffet infantil, e atribua o nome “hipótese alternativa” à afirmação do fornecedor (recorde, o administrador do buffet alega que a proporção de bolas pretas na piscina de bolinhas é 0,3, enquanto que o fornecedor de bolinhas alega que tal proporção é 0,4).

Suponha que a Regra de Decisão continua a mesma que a do Exemplo, porém com a única alteração: agora o conjunto de valores do número de bolas pretas na amostra para os quais a afirmação do administrador será rejeitada é $\{10, 11, \dots, 30\}$ ao invés de $\{11, 12, \dots, 30\}$. Realço que o tamanho de amostra também continua o mesmo: $n = 30$. (Tudo isso pode ser dito da seguinte maneira alternativa: a Regra de Decisão é a mesma mas sua Região Crítica agora é $\{10, 11, \dots, 30\}$.)

(a) Calcule o correspondente Nível de Significância α . Faça esse cálculo usando a distribuição binomial escolhida adequadamente entre as distribuições apresentadas na Figura 9.3.

(b) Suponha que amostra foi feita e acusou 13 bolas pretas entre 30 retiradas. Qual é a conclusão? Na sua formulação da conclusão não esqueça de mencionar o nível de significância

da Regra que produziu a conclusão. Depois, re-escreva a conclusão da maneira que especifica o nível de significância em termos de probabilidade e do evento cuja probabilidade representa-se pelo nível de significância.

Exc. 143. No Exemplo sobre Piscina de Bolinhas, atribua o nome “hipótese nula” à afirmação do fornecedor, e atribua o nome “hipótese alternativa” à afirmação do administrador do buffet infantil.

Suponha agora que o tamanho de amostra e a Regra de Decisão ficaram as mesmas que as do Exemplo, mas a Região Crítica agora é $\{0, 1, \dots, 9\}$ em contraste a $\{1, 2, \dots, 10\}$ que foi analisada no Exemplo.

Comentário. Vale lembrar que a Região Crítica é o conjunto de valores de amostragem que causarão a rejeição da hipótese nomeada “nula”. No âmbito do presente exercício, a Hipótese Nula é a afirmação de fornecedor, enquanto que no Exercício 142, a nula era a afirmação de administrador. Essa diferença causou a diferença entre as Regiões Críticas às quais os exercícios referem.

(a) Calcule o correspondente Nível de Significância α . Faça esse cálculo usando a distribuição binomial escolhida adequadamente entre as distribuições apresentadas na Figura 9.3.

(b) Suponha que amostra foi feita e acusou 13 bolas pretas entre 30 retiradas. Qual é a conclusão? Na sua formulação da conclusão não esqueça de mencionar – da maneira correta – o nível de significância da Regra que produziu a conclusão. Depois, re-escreva a conclusão da maneira que especifica o nível de significância em termos de probabilidade e do evento cuja probabilidade representa-se pelo nível de significância.

Exc. 144. Refaça o item (a) do Exc. 142 para o caso quando $n = 20$ e a Região Crítica é $\{8, 9, \dots, 20\}$. Assim como no Exc. 142, use a distribuição binomial escolhida adequadamente entre as distribuições apresentadas na Figura 9.3. Em sequência, assumo que amostra de 20 bolas apresentou 10 bolas pretas e com esses dados refaça item (b) do Exc. 142.

Exc. 145. Refaça o item (a) do Exc. 143 para o caso quando $n = 20$ e a Região Crítica é $\{0, 1, \dots, 7\}$. Assim como no Exc. 143, use a distribuição binomial escolhida adequadamente entre as distribuições apresentadas na Figura 9.3. Em sequência, assumo que amostra de 20 bolas apresentou 10 bolas pretas e com esses dados refaça item (b) do Exc. 143.

Exc. 146. Refaça o item (a) do Exc. 142 para o caso quando $n = 100$ e a Região Crítica é $\{37, 38, \dots, 100\}$. No seu cálculo do valor de nível de significância use aproximação de binomial pela normal. Em sequência, assumo que amostra de 100 bolas apresentou 29 bolas pretas e com esses dados refaça item (b) do Exc. 142.

Exc. 147. Faça o item (a) do Exc. 143 para o caso quando $n = 100$ e a Região Crítica é $\{0, 1, \dots, 36\}$. No seu cálculo do valor de nível de significância use a aproximação de binomial pela normal. Em sequência, assumo que amostra de 100 bolas apresentou 29 bolas pretas e com esses dados refaça item (b) do Exc. 143.

Exc. 148. No Exemplo sobre Piscina de Bolinhas, atribua o nome “hipótese nula” à afirmação do administrador do buffet infantil, e atribua o nome “hipótese alternativa” à afirmação do fornecedor.

(a) Pedese construir a Região Crítica (RC) da sorte tal que para amostra de tamanho $n = 30$, o correspondente valor do nível de significância (α) esteja o mais próximo possível a 0,05 (ou 5%, se deseja medir em porcentagens); na sua construção, use a distribuição binomial escolhida adequadamente entre as distribuições apresentadas na Figura 9.3. A exigência de ser “mais próximo possível” será explicada no Comentário 2 abaixo.

Comentário 1: Na verdade, espera-se que em todos os exercícios formulados no arcabouço do Exemplo sobre Piscina de Bolinhas, você assuma que a Regra de Decisão siga as seguintes mandamentos. (i) A escolha entre hipóteses nula e alternativa será feita com base na quantidade de bolas pretas na amostra de tamanho $n = 30$. (ii) Se tal quantidade for relativamente pequena, aceitar-se-á a hipótese nula, enquanto que se for relativamente grande, aceitar-se-á a hipótese alternativa. (iii) A quantificação das “pequenas” e “grandes” é assim: há um inteiro denotado por ℓ e chamado de “limiar”; todo que é menor dele considera ser pequeno, tudo que maior ou igual é grande:

$$\underbrace{\{0, 1, \dots, \ell - 1\}}_{\text{valores pequenos}} \quad \underbrace{\{\ell, \ell + 1, \dots, 30\}}_{\text{valores grandes}} \quad (9.34)$$

Dito isso, a Região Crítica adquire o forma $\{\ell, \ell + 1, \dots, 30\}$ no caso do Exc 148, já que na posição da Hipótese Nula foi colocada a afirmação do administrado do buffet infantil, e já que, pela própria definição, a Região Crítica significa o conjunto de tais valores para os quais a Hipótes Nula será rejeitada. Tudo isso junto significa que a parte técnica da tarefa do exercício é achar o valor de ℓ . Mas vale lembrar que isso é válido só se você pensou que a Regra de Decisão foi moldada de acordo com (i)–(iii) supracitados.

Já se você imaginou uma Regra de Decisão que diferencia-se dos moldes (i)–(iii), aí então você e eu teremos problemas graves de comunicação pois toda minha exposição foi feita para corresponder aos (i)–(iii). Portanto, eu sugiro que você largue a “sua” e “incorreta” Regra de Decisão e aceite a minha e a correta, quer dizer, a que atende aos (i)–(iii). Eu não pretendo analisar toda sugestão incorreta que possa ser inventada, pois a quantidade de tais sugestões é enorme devido á quantidade enorme de combinações de atributos nos quais o conjunto de regras (i)–(iii) pode ser alterado. Isso é mais um peso a favor da proibição da liberdade na invenção de regras de decisão e na canonização da única Regra que expressa-se por (i)–(iii).

Antes de fechar o comentário, eu gostaria de mostrar uma das possíveis modificações da Regra de Decisão canonizada. É o assunto do parágrafo abaixo, cuja leitura, alias, pode ser omitida sem qualquer prejuízo para o entendimento da apresentação posterior.

Minha regra modificada coincide com a regra canonizada em tudo, menos a forma das regiões de aceitação/rejeição das hipóteses. Na regra canonizada tais regiões foram determinadas por (9.34). Na regra modificada eu quero que

$$\{0, 1, \dots, \ell - 3, \ell - 1\} \cup \{\ell + 1\} \quad (9.35)$$

sejam os valores para os quais a hipótese “ $p = 0, 3$ ” esteja aceita, enquanto que

$$\{\ell, \ell + 2, \ell + 3, \dots, 29, 30\} \cup \{\ell - 2\} \quad (9.36)$$

sejam os valores para os quais a hipótese “ $p = 0, 4$ ” seja aceita. (Espero que você percebeu a idéia por trás da construção (9.35)-(9.36): retira-se o valor $\ell - 2$ do primeiro conjunto e leve-se para o segundo, mas, para compensar, retira-se $\ell + 1$ do segundo e leva-se para o primeiro.) Se a divisão (9.35)-(9.36) fosse de seu agrado, eu não teria como lhe despersuadir com um argumento simples e sólido. A única ferramenta de despersuasão que possui é o princípio de Máxima Verossimilhança. Mas, honestamente falando, ela me permitiria indicar que a divisão (9.34) é razoável mas não permitiria provar que a divisão (9.35)-(9.36) é inadequada. Esse pequeno exemplo corrobora minha tese de que a Regra de Decisão baseada em (i)–(iii) deve ser canonizada sem discussões e desafios.

Comentário 2 (sobre o sentido da exigência de ser “mais próximo possível” na formulação da tarefa do exercício). Segindo o acordo firmado no Comentário 1 acima, nos exercícios formulados no arcabouço do Exemplo sobre Piscina de Bolinhas, o nível de significância será sempre igual à uma das duas seguintes probabilidades

$$IP[X \in \{0, 1, \dots, \ell - 1\}], \quad IP[X \in \{\ell, \ell + 1, \dots, 30\}], \quad (9.37)$$

onde a variável aleatória X tem distribuição binomial

Acontece que pode não existir o número natural ℓ tal que a segunda probabilidade seja igual exatamente a 0,05 (assumindo que a probabilidade que nos interessa é a segunda das duas em (9.37)). Entretanto, sempre existe um número natural ℓ para o qual a probabilidade é a mais próxima ao desejado 0,05. Infelizmente, podem existir dois valores que dão a mesma proximidade: uma por de cima e outra por de baixo de 0,05. Nesse caso, vamos concordar que seja escolhido aquele ℓ cuja probabilidade é maior que 0,05.

Item (b) do Exc. 148. Suponha que amostra foi feita e acusou 13 bolas pretas entre 30 retiradas. Qual é a conclusão da Regra de Decisão com a Região Crítica construída no item (a)? Na sua formulação da conclusão não esqueça de mencionar o nível de significância da Regra que produziu a conclusão. Depois, re-escreva a conclusão da maneira que especifica o nível de significância em termos de probabilidade e do evento cuja probabilidade representa-se pelo nível de significância.

Exc. 149. No Exemplo sobre Piscina de Bolinhas, atribua o nome “hipótese nula” à afirmação do fornecedor, e atribua o nome “hipótese alternativa” à afirmação do administrador do buffet infantil.

(a) Pedese construir a Região Crítica da sorte tal que para amostra de tamanho $n = 30$, o correspondente valor do nível de significância (α) esteja o mais próximo possível a 0,1 (ou 10%, se deseja medir em porcentagens).

(b) Suponha que amostra foi feita e acusou 13 bolas pretas entre 30 retiradas. Qual é a conclusão da Regra de Decisão com a Região Crítica construída no item (a)? Na sua formulação da conclusão não esqueça de mencionar o nível de significância da Regra que produziu a conclusão. Depois, re-escreva a conclusão da maneira que especifica o nível de significância em termos de probabilidade e do evento cuja probabilidade representa-se pelo nível de significância.

Exc. 150. Refaça o item (a) do Exc. 148 para o caso quando $n = 20$; no seu cálculo da Região Crítica use a distribuição binomial escolhida adequadamente entre as distribuições apresentadas na Figura 9.3. Em sequência, assuma que amostra de 20 bolas apresentou 7 bolas pretas e com esses dados refaça item (b) do Exc. 148.

Exc. 151. Refaça o item (a) do Exc. 149 para o caso quando $n = 20$; no seu cálculo da Região Crítica use a distribuição binomial escolhida adequadamente entre as distribuições apresentadas na Figura 9.3. Em sequência, assuma que amostra de 20 bolas apresentou 10 bolas pretas e com esses dados refaça item (b) do Exc. 149.

Exc. 152. Refaça o item (a) do Exc. 148 para o caso quando $n = 100$; no seu cálculo da Região Crítica use aproximação de binomial pela normal. Assuma em sequência que amostra de 100 bolas apresentou 29 bolas pretas e com esses dados refaça item (b) do Exc. 148.

Exc. 153. Refaça o item (a) do Exc. 143 para o caso quando $n = 100$; no seu cálculo da Região Crítica use aproximação de binomial pela normal. Assuma em sequência que amostra de 100 bolas apresentou 29 bolas pretas e com esses dados refaça item (b) do Exc. 149.

9.11.2 Exercícios sobre os testes nos quais um só ou nenhuma hipótese é precisa

Exc. 154. Suponha que as drogas usuais para leucemia provoquem efeitos colaterais em 60% dos pacientes. Um laboratório consegue eliminar de certo medicamento um radical acetil, com isso considera estar diante de uma nova droga com o mesmo poder de cura, mas com menor atividade indesejável. O laboratório espera, portanto, que a proporção de indivíduos com efeitos

adversos tratados com essa nova droga seja menor do que com as drogas usuais. O laboratório resolve testar essa afirmação, aplicando a nova droga a alguns pacientes. Assume, para o efeito de formulação das hipóteses, que sabe-se por certo que a modificação feita no medicamento pelo laboratório não pode aumentar a proporção dos seus usuários nos quais provoca efeitos colaterais.

(a) Formule este problema como um problema de testes de hipóteses.

(b) O laboratório conseguirá testar a nova droga em 19 pacientes. Para este tamanho de amostra, construa a região crítica correspondente ao nível de significância igual a 9%. Em sua construção, use diretamente a distribuição Binomial (quer dizer, não empregue a aproximação da Binomial pela Normal em sua solução). Devido ao uso direto da distribuição Binomial pode ocorrer que seja impossível alcançar o valor exato de 9% para o nível de significância; se for o caso, procure pelo nível mais próximo ao 9% mas não superior a este. Use a região crítica construída para dar resposta correspondente ao resultado “9 das 19 pessoas testadas apresentaram o efeito colateral”.

(c) *Esse item não pode e não deve ser visto como uma continuação do tratamento do exercício proferido nas respostas aos itens (a) e (b). Explicitamente falando, o presente item relaciona-se ao enunciado e ao item (a), mas não tem nada a ver com item (b).*

A afirmação do laboratório será testado em 200 pacientes, que receberão a nova droga, e será aceito caso não mais que 118 dos testados apresentem efeitos adversos. Identifique a Região Crítica e calcule o nível de significância (no cálculo do nível de significância, use a aproximação de binomial por normal). Dé a resposta, caso 119 dos pacientes testados apresentem efeitos adversos.

Exc. 155. Para verificar se uma moeda é honesta, com base em 24 lançamentos independentes, adotamos o seguinte critério: consideramos a moeda não honesta se o resultado for menor do que 8 ou maior do que 16.

(a) Formule esse problema como um problema de teste de hipóteses, descrevendo a regra de decisão e a região crítica deste teste.

(b) Qual é o nível de significância do teste?

(c) Qual é a probabilidade de concluir que a moeda é honesta se a probabilidade de sair cara for 0,3?

Exc. 156. É conhecido por certo que, anualmente, exatamente 20% dos alunos da disciplina “Estatística Básica” eram reprovados. A partir do ano 2004 foi usado um novo método de ensino desta disciplina que é mais amplo que o método que era usado antes. Portanto, a porcentagem dos reprovados entre os alunos, que aprendem por novo método, não pode ser superior a 20%. Queremos identificar se este novo método reduz a porcentagem da reprovação. Para isso serão monitorados 100 alunos que estudam a disciplina no 1-o semestre de 2004. A escolha entre as hipóteses

H_0 : a porcentagem p de reprovação por novo método é exatamente 20%

H_A : a porcentagem p de reprovação por novo método é menor que 20%

será feita com base na proporção de alunos monitorados que serão reprovados na final da disciplina. Queremos que o nível de significância do teste seja 5%. Qual é a Região Crítica do teste? Qual será a resposta do teste, caso a proporção amostral dos reprovados for 18%?

Exc. 157. (Obs.: *Este exercício é muito parecido com Exc. 156, mas com “maior” no lugar do “menor” em sua hipótese alternativa. Preste a atenção como esta substituição afeta o sinal $>$ na sua solução (exatamente, na conta do nível de significância). Esta observação pode te ajudar a entender melhor a escolha entre “ $<$ ” e “ $>$ ” na conta do valor de nível de significância em todos os exercícios do tema.*)

Sabe-se que 70% dos alunos da disciplina “Estatística Básica” são aprovados na disciplina. A

partir do ano 2004 foi usado um novo método de ensino desta disciplina que é mais amplo que o método que era usado antes. Portanto, a porcentagem dos aprovados entre os alunos, que aprendem por novo método, não pode ser inferior a 70%. Queremos identificar se este novo método aumenta a porcentagem da aprovação. Para isso serão monitorados 100 alunos que cursam a disciplina no 1º semestre de 2012.

(a) A escolha entre as hipóteses

H : a porcentagem p de aprovação por novo método é exatamente 70%

A : a porcentagem p de aprovação por novo método é maior que 70%

será feita com base no número de alunos aprovados daqueles 100 alunos da amostra (quer dizer, os monitorados): se o número de aprovados for 73 ou mais, será decidido que o novométodo de ensino é melhor de que o antigo. Qual é a região crítica e qual é o valor do nível de significância desse procedimento?

(b) (*Esse item é independente do item (a); só a formulação das hipóteses é a mesma.*) A escolha entre H e A será feita com base na proporção de alunos monitorados que serão aprovados no final da disciplina. Queremos que o nível de significância do teste seja 10%. Qual é a Região Crítica do teste? Qual será a conclusão do teste, caso a proporção amostral dos aprovados seja de 74%?

Exc. 158. Dez por cento dos veículos que atravessam uma fronteira levam contrabando. As autoridades alfandegárias e de fronteira elaboram um plano para coibir o contrabando e garantem que essa porcentagem diminuiu. Inspetores testam a afirmação examinando uma amostra de 200 veículos após o plano ser colocado em prática; eles concluirão que o plano esteja bem sucedido caso o exeme revelar 15 ou menos veículos com contrabando.

Formule o teste de hipótese estatística adequado e calcule seu nível de significância.

Exc. 159. A marca Z de um produto é responsável por 50% das vendas desse produto em um supermercado. Uma campanha promocional foi contratada e os promotores garantem que a marca Z passará a ser responsável por uma porcentagem maior das vendas. O dono do supermercado propõe entrevistar 18 clientes após o encerramento da campanha promocional e perguntar a cada um deles se compra a marca Z do produto. Sendo p a porcentagem de vendas do produto Z após a campanha, decide-se rejeitar $H : p = 0,50$ em favor de $A : p > 0,50$, se dentre os 18 entrevistados pelo menos 14 deles comprarem a marca Z do produto.

(a) Qual é o significado do erro do tipo I e do tipo II para o problema?

(b) Qual é a região crítica e o nível de significância do teste?

(c) Se a proporção p de clientes que compram a marca Z do produto realmente aumentar para 80% após a campanha, calcule a probabilidade dessa alteração não ser detectada.

Exc. 160. Sabe-se através de experiências passadas que, se uma determinada máquina estiver ajustada, apenas 5% dos itens por ela produzidos serão defeituosos. Diariamente são inspecionados os primeiros 25 itens produzidos pela máquina. Se o número de itens defeituosos for, no máximo 3, a produção continua sem interrupção. Caso sejam selecionados 4 ou mais itens defeituosos pára-se a máquina para que ela seja ajustada.

a) Formule este problema como um problema de teste de hipóteses especificando as hipóteses nula e alternativa.

b) Quais são os significados práticos dos erros tipo I e tipo II?

c) Qual é a região crítica do teste?

d) Qual é o nível de significância do teste?

e) Agora esqueça da regra até 3 peças defeituosas em 25 testadas, e reponda a seguinte pergunta: O responsável pela produção dispõe a testar 50 peças mas gostaria de ter, no máximo, uma chance em 100 de parar desnecessariamente a produção para ajustar a máquina. Qual deveria ser uma região crítica para alcançar esse objetivo?

Exc. 161. Um relatório de uma companhia afirma que 40% de toda a água obtida através de poços artesianos no Nordeste é salobra. Há muitas controvérsias sobre essa informação, alguns dizem que a proporção é maior, outros que é menor. Para esclarecer a dúvida, 30 poços foram sorteados e a afirmação do relatório foi considerada falsa se o número de poços com água salobra for menor do que 8 ou maior do 16.

- Formule esse problema como um problema de teste de hipóteses.
- Interprete os erros tipo I e tipo II.
- Qual é a região crítica do teste?
- Qual é o nível de significância do teste?

Exc. 162. O Departamento de Esportes de um determinado clube de São Paulo está selecionando jogadores de basquete para o próximo campeonato paulista. Um dos testes aplicados, que é eliminatório, consiste em solicitar ao jogador que arremesse 100 bolas da posição de arremessos livres. O clube pretende selecionar os atletas com aproveitamento superior a 50% e para isso adota como a regra de decisão aprovar todo atleta que converter pelo menos 60 cestas.

- Formule este problema como um problema de testes de hipóteses especificando as hipóteses nula e alternativa. Qual a região crítica adotada no teste? (Observação: nula $\leq 50\%$, alternativa $> 50\%$.)
- Descreva em palavras os erros de tipo I e de tipo II. Determine a probabilidade de erro de tipo I. Qual o nível de significância do teste?
- Determine a probabilidade de que o clube deixe de selecionar um atleta cujo aproveitamento seja na verdade de 70%.

Exc. 163. Suspeitamos que a moeda *D. Pedro V* de 200 Réis (do ano 1855) seja não honesta, pois o relevo da coroa num dos seus lados faz com que tal lado seja mais pesado e, consequentemente, a probabilidade da moeda cair com “cara” para cima deva ser maior que a de cair com “coroa” para cima. Então, para testar as hipóteses

$$H: p = 1/2$$

contra a hipótese

$$A: p > 1/2,$$

onde p denota a probabilidade da moeda mostrar “cara” num lançamento, vamos lançar a moeda 120 vezes, e caso observar 64 ‘caras ou mais, decidiremos a favor da hipótese A ; caso contrários, decidiremos a favor de H . Pede-se calcular o nível de significância desse teste.

9.11.3 Exercício sobre teste de hipóteses sobre média

Exercício 164. REQUER VERIFICAÇÃO. A diretoria do fabricante do refrigerante "fresca-cola" resolveu que retiraria o produto da linha de produção se a média de consumo mensal per capita fosse menor que 8 litros, caso contrário continuaria a fabricá-lo. Sabe-se que o desvio padrão da produção mensal é 2,2 litros. Foi realizada uma pesquisa de mercado, tomando-se uma amostra aleatória de 80 indivíduos e verificou-se que o consumo mensal médio do refrigerante para eles foi de 7,4 litros. Formule o problema como um teste de hipóteses estatístico e conclua, com base no nível descritivo, se a diretoria deve ou não retirar o "fresca-cola" do mercado, ao nível de significância de 5%. Justifique sua resposta.

Exercício 165. REQUER VERIFICAÇÃO. A FIESP quer avaliar o tempo ocioso em consequência de acidentes de trabalho nesta década. Na década de 2000, a média de tal tempo era de 60 horas/homem/ano. Ela então patrocinou um estudo em que se tomou uma amostra de 100 indústrias e observou-se que o número médio de horas/homem perdidas por acidente ano passado foi 50 horas com um desvio padrão de 20 horas/homem.

- Formule o problema como um teste de hipóteses estatístico.
- Qual o significado dos erros de tipo I e tipo II no problema.
- Determine uma região crítica para um nível de significância de 10
- Com base na região crítica construída em (c), você diria que houve redução no tempo perdido em acidentes de trabalho comparado com a década passada? Justifique sua resposta.

Exercício 166. REQUER VERIFICAÇÃO. É conhecido por certo, que o peso de bois de uma fazenda segue distribuição normal com média igual a 200 kg e desvio padrão igual a 10 kg. Os bois começaram a receber um alimento suplementar, que não pode diminuir o peso. Para verificar que este alimento causa um aumento de peso, iremos colher uma amostra de 400 bois e escolher entre as hipóteses

H_0 : o peso da população tem distribuição $\mathcal{N}(200; 10^2)$;

H_A : o peso da população tem distribuição $\mathcal{N}(\mu; 10^2)$ com $\mu > 200$

com base no peso médio dos bois da amostra. É desejado que o nível de significância do teste seja 7%. Qual é a Região Crítica do teste? Qual será a resposta do teste, caso o peso média amostral for 210 kg?

Exercício 167. REQUER VERIFICAÇÃO. É conhecido por certo, que o tempo de vida de galinhas de uma certa fazenda segue distribuição normal com média igual a 10 anos e desvio padrão igual a 1 ano. As galinhas começaram a receber um alimento suplementar que não pode diminuir o tempo da vida. Para verificar que este alimento causa um aumento do tempo de vida, iremos observar 400 galinhas da fazenda e escolher entre as hipóteses

H_0 : o tempo de vida da população tem distribuição $\mathcal{N}(10; 12)$; H_A : o tempo de vida da população tem distribuição $\mathcal{N}(\mu; 12)$ com $\mu > 10$

com base no tempo médio da vida das galinhas da amostra. É desejado que o nível de significância do teste seja 8%. Qual é a Região Crítica do teste? Qual será a resposta do teste, caso o tempo de vida médio amostral seja 10,5 anos?

Exercício 168. REQUER VERIFICAÇÃO. Sabe-se que uma máquina de encher pacotes de café coloca pó de café conforme uma distribuição normal com o desvio padrão 0,05 kg. A máquina foi regulada para que o peso médio de pacotes de café seja 0,5 kg. Desconfia-se que ela se desregulou e coloca pó em quantidade média menor que 0,5 kg. Para escolher entre

H_0 : o peso dos pacotes produzidos pela máquina tem distribuição $\mathcal{N}(0,5; (0,05)^2)$;

H_A : o peso dos pacotes produzidos pela máquina tem distribuição $\mathcal{N}(\mu; (0,05)^2)$ com $\mu < 0,5$ colheremos 100 pacotes da produção e faremos o peso médio deles. É desejado que o nível de

significância do teste seja 3%. Qual é a Região Crítica do teste? Qual será a resposta do teste, caso o peso médio de 100 pacotes seja 4,98?

Exercício 169. REQUER VERIFICAÇÃO. Em um estudo sobre o salário de empregadas domésticas na cidade de São Paulo, foram sorteadas e entrevistadas 200 trabalhadoras. Sabe-se que o desvio padrão desse salário 0,8 mínimos. A amostra forneceu média de 2,5 salários mínimos. Deseja-se testar se a média populacional dos salários das empregadas domésticas na cidade de São Paulo é igual a 3 salários mínimos ou é menor.

- Formule o problema como um teste de hipóteses estatístico.
- Determine uma região crítica para um nível de significância de 5%.
- Com base na região crítica construída em (b), qual é a sua conclusão? Justifique a resposta.

Exercício 170. REQUER VERIFICAÇÃO. Um fabricante de lâmpadas afirma que seu produto tem duração média de 400 horas. Uma amostra de 25 lâmpadas forneceu vida média amostral igual há 395 horas. Se o desvio padrão do tempo de vida das lâmpadas é igual há 30 horas,

- formule o problema como um teste de hipóteses estatístico.
- Determine a região crítica para um nível de significância de 0,05
- Com base nessa região crítica, qual é a conclusão?
- Determine a probabilidade do erro do tipo II, se o tempo médio de vida for 385 h.

Exercício 171. (Ex.10.1 do livro “Estatística básica” do P. A. Morettin e W. O. Bussab (página 234).) REQUER VERIFICAÇÃO.

Teste sobre a média de uma população, com variância conhecida.

Uma indústria usa como um dos componentes das máquinas que produz um parafuso especial importado, que deve satisfazer a alguns exigências. Uma destas exigências é a resistência à tração. Esses parafusos são fabricados por dois países: Estados Unidos e Japão. Nos ambos casos, a resistência dos parafusos à tração segue a distribuição normal. Para os americanos, a média desta distribuição é de 145 kg, e o desvio padrão é de 12 kg. Já para parafusos japoneses, a média é de 155 kg e desvio padrão é de 20 kg.

Um lote destes parafusos, de origem desconhecida, será leiloado a um preço muito convidativo. O edital do leiloeiro afirma que, pouco antes de leilão, será divulgada a resistência média de uma amostra de 25 parafusos do lote. O representante da indústria no leilão resolve concluir que o lote contém parafusos japoneses se a média amostral acusar um valor > 150 kg, e concluir que são parafusos americanos. Em caso que a conclusão dele seja “os parafusos são japoneses”, ele comprará o lote.

- Formule este problema como um problema de testes de hipóteses especificando as hipóteses nula e alternativa. Qual a região crítica adotada no teste?
- Descreva em palavras os erros de tipo I e de tipo II. Determine as probabilidades de erro de tipo I e de erro do tipo II. Qual o nível de significância do teste?

9.12 Soluções dos exercícios sobre Teste de Hipóteses para Proporção

9.12.1 Soluções dos exercícios da Seção 9.11.1

Solução do Exc. 142(a). O raciocínio que leva ao cálculo do nível de significância profere-se “no ambiente alegado pela Hipótese Nula”. A execução técnica disso segue o seguinte mandamento: assumir que a hipótese nula é verdadeira e calcular a probabilidade da Regra de Decisão rejeitar a verdade, quer dizer, rejeitar a Hipótese Nula. Tal probabilidade é, pela própria definição, o nível de significância do teste. Ao executar tal cálculo, é preciso ter em mente que sua abordagem não significa que a realidade coincide com a alegação da Hipótese Nula. Para que isso esteja claro, eu vou usar frase do tipo “qual seria a probabilidade de erro se a Hipótese Nula fosse verdade”, e também vou chamar por “realidade virtual” o ambiente no qual proferem-se os cálculos.

Um outro princípio que ampara o cálculo do nível de significância é que tal cálculo faz-se imaginando que a amostra ainda não tinha sido proferida. Desse ponto de vista, o resultado de amostragem torna-se uma variável aleatória.

Então, assumimos a realidade virtual alegada pela Hipótese Nula. Nessa realidade, a quantidade de bolas pretas a serem vistas na amostra de tamanho $n = 30$ segue a distribuição $\text{Bin}(30, 0, 3)$. (Realço: “0, 3” surgiu na distribuição pois há 30% de bolas pretas na piscina, de acordo com a realidade virtual assumida.) Para a comodidade na exposição, denotamos por X a variável aleatória com a distribuição $\text{Bin}(30, 0, 3)$. Recordamos que a Regra de Decisão irão dizer que a proporção de bolas pretas na piscina não é 30% se X assumir valores na Região Crítica da Regra; tal região é $\{10, 11, \dots, 30\}$ no presente exercício.

- ▷ Na verdade, sabemos que a Regra de Decisão não somente rejeita que “ $p = 0, 3$ ” mas afirma que “ $p = 0, 4$ ”, quer dizer, junto com a rejeição da Hipótese Nula vem a afirmação precisa sobre o valor alternativo de p que está válido na visão da Regra. Tal precisão é impossível quando a Hipótese Alternativa não é precisa. Isso é o caso muito comum. Portanto, com o intuito de dar a maior generalidade ao presente argumento, eu escrevi acima “a Regra rejeitará que $p = 0, 3$ ” ao invés de “a Regra aceitará que $p = 0, 4$ ”.

Portanto,

$$\mathbb{P}[X \in \{10, 11, \dots, 30\}]$$

é a probabilidade da Regra errar no ambiente da realidade virtual alegada pela Hipótese Nula. Juntando as peças, concluímos que nos interessa é

$$\mathbb{P}[X \in \{10, 11, \dots, 30\}], \text{ onde } X \sim \text{Bin}(30, 0, 3) \quad (9.38)$$

Para calcular o valor da probabilidade em interesse, usamos a primeira coluna da Tabela na Figura 9.3. A resposta seria o valor da seguinte somatória

$$\mathbb{P}[X = 10] + \mathbb{P}[X = 11] + \dots + \mathbb{P}[X = 30]$$

Mas o mesmo valor é

$$1 - (\mathbb{P}[X = 0] + \mathbb{P}[X = 1] + \dots + \mathbb{P}[X = 9])$$

enquanto que seu cálculo é mais curto. Então, a resposta, quer dizer, o nível de significância perguntado é

$$\begin{aligned} 1 &- (0,000023 + 0,0003 + 0,0018 + 0,0072 + 0,0208 + 0,0464 + 0,083 + 0,122 + 0,150 + 0,157) \\ &= 0,411477 \approx 0,41 \end{aligned}$$

O valor é meio alto para os padrões de aplicação, mas isso não deve nos chatear, pois nosso objetivo aqui é puramente teórico.

Solução do Exc. 142(b). Já que 13 está na Região Crítica da Regra de Decisão, então – devido a própria definição da Região Crítica – devemos rejeitar a Hipótese Nula. A resposta formal soa assim: “Ao nível de significância de 0,41, com a observação de 13 bolas pretas na amostra de 30, deve-se rejeitar a Hipótese Nula, isto é, deve-se rejeitar que a proporção de bolas pretas na urna é 0,3”.

A observação “13 bolas” na amostra de tamanho 30 caiu na região definido como a região de rejeição da Hipótese Nula, a que afirma que a proporção de bolas pretas na população é 0,3. Entretanto, na validade da Hipótese Nula, as observações poderiam assumir valores da Região Crítica; isso aconteceria com a probabilidade 0,411 (que é o nível de significância correspondente à Região Crítica). Portanto, a Regra de Decisão aplicada rejeita com a probabilidade 0,411 a Hipótese Nula quando essa é verdadeira. Na prática, isso significa que a observação “13” poderia acontecer sob a validade da Hipótese Nula, embora tal observação foi a razão para a rejeição dessa hipótese.

Solução do Exc. 143(a). Como a Hipótese Nula agora é a hipótese que afirma que $p = 0,4$, então a realidade virtual para a qual faremos as contas é que a proporção de bolas pretas na piscina é 0,4. Ao representar por variável aleatória X a quantidade de bolas pretas a ser vista na amostra a ser retirada (a amostra é de tamanho $n = 30$), temos que $X \sim Bin(30, 0,4)$. Ao juntar isso com o formato da Região Crítica definido pelo enunciado do presente exercício, conclui-se que o correspondente nível de significância é

$$\mathbb{P}[X \in \{0, 1, \dots, 9\}], \text{ onde } X \sim Bin(30, 0,4) \quad (9.39)$$

O cálculo desse valor usa a segunda coluna da tabela apresentada na Figura 9.3.

Comentário I acrescido em 03/12. Em soluções de exercícios sobre teste de hipóteses as fórmulas (9.38) e (9.39) são apresentadas da seguinte maneira alternativa:

$$\mathbb{P}[X \geq 10] \text{ ao envez de } \mathbb{P}[X \in \{10, 11, \dots, 30\}], \mathbb{P}[X \leq 9] \text{ ao envez de } \mathbb{P}[X \in \{1, 2, \dots, 9\}] \quad (9.40)$$

É óbvio que os sinais “ \geq ” e “ \leq ” nas probabilidades da primeira coluna não podem ser trocadas. Apesar disso, o uso de desigualdade errada é o erro mais frequente que vi nas provas de alunos da disciplina Estatística Básica e afins. Eu não identifiquei a razão objetiva por trás desse erro e portanto não consigo criar exercícios auxiliares que lhe ajudariam a se vacinar contra o erro. Só posso lhe pedir a atenção no uso de desigualdade.

Solução do Exc. 143(b).

A solução do Exc. 144 é parecida com a do Exc. 142. A diferença principal é que agora $n = 20$, fato que faz com que a distribuição de X agora é $Bin(20, 0,3)$.

A solução do Exc. 145 é parecida com a do Exc. 143. A diferença principal é que agora $n = 20$, fato que faz com que a distribuição de X agora é $Bin(20, 0,4)$.

A solução do Exc. 146 é parecida com a do Exc. 142. A diferença principal é que agora $n = 100$, fato que faz com que a distribuição de X agora é $Bin(100, 0,3)$. Para valor tal alto de n é cômodo usar a aproximação de Binomial pela normal. Sabe-se que variável aleatória normal Y será boa aproximação para X se a média de Y coincidir com a média de X e a variância de Y coincidir com a variância de X . Já que X é binomial, sabemos que $\mathbb{E}[X] = 100 \times 0,3 = 30$ e $\text{Var}[X] = 100 \times 0,3(1 - 0,3) = 21$. Portanto o valor da nível de significância, cuja expressão exata é

$$\mathbb{P}[X \in \{37, 38, \dots, 100\}], \text{ onde } X \sim Bin(100, 0,3)$$

aproxima-se pela

$$\mathbb{P}[Y \geq 37], \text{ onde } Y \sim \mathcal{N}(30, 21)$$

Solução do Exc. 147. A ser colocada.

Solução do Exc. 148(a). Como $H : p = 0,3$ e $A : p = 0,4$ então a divisão entre os possíveis valores de X (o número de bolas pretas na amostra) deve seguir a regra: observa-se valor grande de $X \Rightarrow$ aceita-se A ; observa-se valor pequeno de $X \Rightarrow$ aceita-se H . Portanto, a Região Crítica deve ter o formato $\ell, \ell + 1, \dots, 30$. Com todas estas definições, o nível de significância (α) é

$$\mathbb{P}[X \in \{\ell, \ell + 1, \dots, 30\}], \text{ onde } X \sim \text{Bin}(30; 0,3)$$

Por exemplo, se $\ell = 19$, então, de acordo com os valores da primeira coluna da tabela na Figura 9.3, tem-se:

$$\begin{aligned} \alpha &= \mathbb{P}[X = 19] + \mathbb{P}[X = 20] + \dots + \mathbb{P}[X = 30] \\ &= 0,000125 + 0,00003 + 0,000006 + 0,000001 + 0,00000016 + \text{valores insignificantes pequenos} \\ &= 0,00016216 \end{aligned}$$

O resultado não deu 5%, e é por isto que a Região Crítica no caso não é $\{19, 20, \dots, 30\}$. A tarefa é achar ℓ para que o resultado da conta do tipo acima desse 0,05 (5%, quer dizer).

9.12.2 Soluções dos exercícios da Seção 9.11.2

Solução do Exc. 154. Resposta em (a). Denotado por p a proporção populacional das pessoas para as quais a nova droga provoca efeitos colaterais, temos que o problema do enunciado pode ser reformulado como o problema de escolha entre

H (a hipótese nula): $p = 0,6$ e

A (a hipótese alternativa): $p < 0,6$.

Observe que o rótulo “nula” recebeu aquela das duas hipóteses que tem a afirmação precisa, isto é, aquela que afirma a igualdade.

Observe também que a distribuição dos nomes “nula” e “alternativa” por hipóteses não tem nada ver com que uma das hipóteses seja “positiva” ou “desejada” enquanto que a outra seja a obra do advogado do diabo. No presente caso, por exemplo, todos nós desejamos que o novo remédio tenha eficácia elevada, quer dizer, todos queremos que a hipótese $p < 0,6$ vença. Outrassim, essa hipótese adquiriu o nome “alternativa”.

Resposta em (b). Para resolver item (b), precisamos fingir que a realidade corresponde à afirmação da hipótese nula e nos colocar no eixo de tempo antes da aplicação experimental da droga num grupo de pacientes com o objetivo de estudar os efeitos colaterais. De acordo com o enunciado do item (b), este grupo tem 19 pacientes. Do ponto de vista da posição de “antes” posição, o número de pacientes dentre dos 19 a serem testados, que apresentarão efeito colateral, é uma variável aleatória. Denotamos esta por X . Como “fingimos” que a hipótese nula corresponde à realidade, temos portanto que

$$X \text{ tem distribuição binomial com } n = 19 \text{ e } p = 0,6$$

A determinação da distribuição da quantidade a ser observada (é a que foi denominada acima por X) é um dos ingredientes da solução. O segundo ingrediente é a determinação da região crítica. Recordamos que “região crítica” é nome para o conjunto dos valores de X para os quais a regra rejeita a hipótese nula. A regra da construção deste conjunto está embutida no enunciado de exercício e revela-se para nós logo que formulamos as hipóteses. Esta é o

princípio genérico. Eis como ele realiza-se no presente caso. As hipóteses discordam entre si sobre o valor de p e esta discordância pode ser resumida verbalmente da seguinte maneira: a hipótese alternativa afirma que o valor de p é menor do que aquele afirmado pela hipótese nula. Formularemos agora a repercussão desta discordância na distribuição de X : se a verdade for a da afirmação da hipótese nula, então a distribuição de X é aquela descrita acima, enquanto que se a verdade for da hipótese alternativa então o parâmetro p desta distribuição deve ser menor que 0,6 e, conseqüentemente, X deve assumir valores pequenos com probabilidades maiores. Esta consequência é o fundamento para a regra da decisão a ser tomada com base na observação do valor de X : se a valor for pequena - rejeitar a hipótese nula, caso contrário - aceitar esta hipótese. A mesma coisa pode ser dita com o emprego da informação de que X assume valores em $\{0, 1, 2, \dots, 19\}$ seja para $p = 0,6$ ou seja para $p > 0,6$: se X assumir valor no conjunto $\{0, 1, 2, \dots, k\}$ - rejeitar a hipótese nula, já se for do conjunto $\{k + 1, k + 2, \dots, 19\}$ - aceitar a hipótese nula. É evidente que ainda não dissemos nada sobre o valor de k que apareceu nos conjuntos. Sobre este versaremos daqui a pouco. Agora só concluímos a resposta na pergunta de identificação da região crítica: este é o conjunto $\{0, 1, 2, \dots, k\}$. Denotaremos este por \mathcal{RC} , isto é:

$$\mathcal{RC} = \{0, 1, \dots, k\}$$

O valor de k , extremo direito da \mathcal{RC} , virá do valor que desejamos a ser atingido pelo nível de significância de nosso teste (recorde, o nível de significância denota-se por α). A relação deste com k é dada via

$$\alpha = \mathbb{P}[X \in \mathcal{RC}] = \mathbb{P}[X \leq k],$$

onde – vale lembrar – a distribuição de X é a que foi construída com base no pressuposto da validade da afirmação da hipótese nula. Ela é $\text{Bin}(19; 0,6)$. Eis a tabela dela:

0	2.748779e-08	10	1.464269e-01
1	7.834020e-07	11	1.797058e-01
2	1.057593e-05	12	1.797058e-01
3	8.989538e-05	13	1.451470e-01
4	5.393723e-04	14	9.330877e-02
5	2.427175e-03	15	4.665439e-02
6	8.495114e-03	16	1.749540e-02
7	2.366496e-02	17	4.631134e-03
8	5.324616e-02	18	7.718557e-04
9	9.761796e-02	19	6.093597e-05

Usando a tabela, precisamos achar k que corresponde a valor 9% de α , ou, se tal valor não existir, ao valor mais próximo a 9%. Somando os valores da tabela, tem-se que

$$\begin{aligned} \mathbb{P}[X \leq 7] &= 0,0352279 \\ \mathbb{P}[X \leq 8] &= 0,08847406 \\ \mathbb{P}[X \leq 9] &= 0,1860920 \\ \mathbb{P}[X \leq 10] &= 0,3325190 \end{aligned}$$

Estes valores indicam que $k = 8$, quer dizer a região crítica é $\{0, 1, \dots, 7, 8\}$.

Resposta em (c) Tem-se diretamente do enunciado do item que

$$R_C = \{0, 1, \dots, 117, 118\}$$

Para calcular o nível de significância (α) é precisa entender que, do ponto de vista de antes da contagem do número dos pacientes, que apresentam efeitos colaterais, tal número pode ser visto como uma variável aleatória (a ser denotada por X) e que essa tem a distribuição $\text{Bin}(200; 0,6)$

caso a hipótese nula for verdade. Ressalvo que não sabemos por certo qual das duas hipóteses é verdadeira, mas o cálculo do nível de significância calcula-se sob a suposição da veracidade da hipótese nula, e essa, no caso é “ $p = 0,6$ ”. Como, pela própria definição, o nível de significância é a probabilidade de rejeitar a hipótese nula quando ela é verdadeira, tem-se que

$$\alpha = \mathbb{P}[X \in R_C] = \mathbb{P}[X \leq 118] \text{ onde } X \sim \text{Bin}(200; 0,6)$$

Como $\mathbb{E}[X] = 200 \times 0,6 = 120$, $\text{Var}[X] = 200 \times 0,6 \times (1 - 0,6) = 48$, então X pode ser aproximada pela $Y \sim \mathcal{N}(120; (6,93)^2)$, aproximação essa que dá:

$$\mathbb{P}[X \leq 118] \approx \mathbb{P}[Y \leq 118] = \mathbb{P}\left[Z \leq \frac{118 - 120}{6,93}\right] = \mathbb{P}[Z \leq -0,29] = 1 - 0,6141 \approx 0,39$$

Com o resultado “119 pessoas das 200 tratadas com nova droga apresentaram efeitos colaterais” a hipótese nula não será rejeitada pois 119 está fora da R_C . Isso significa em termos do enunciado, que o teste cujo nível de significância foi 0,39 não aceitou a afirmação do laboratório que sua nova droga diminui o efeito colateral para mais baixo que 60% dos pacientes que a tomam.

Solução do Exc. 155. Resposta em (a). Denotamos por p a probabilidade da moeda considerada a dar “cara” num lançamento. Responderemos à questão com uso desta notação. Isto nos permitirá aproveitar esta resposta nas respostas aos itens (b) e (c).

Segundo o acordo de nomenclatura, uma moeda chama-se honesta se $p = 0,5$. Portanto, a dúvida de se nossa moeda seja honesta ou não pode ser expressa como a tarefa de escolha entre as duas seguintes hipóteses:

$p = 0,5$, afirmação a ser rotulada “hipótese nula” e denotada por H ; e:

$p \neq 0,5$, afirmação a ser rotulada “hipótese alternativa” e denotada por A .

Quanto à tarefa de descrição da regra de decisão, notamos que tal regra foi formulada no enunciado da seguinte maneira: “adotamos o seguinte critério: consideramos a moeda não honesta se o resultado for menor do que 8 ou maior do que 16”. Esta é a resposta na tarefa.

Quanto à descrição da região crítica, começaremos com o lembrete de que esta é o conjunto das observações que nós levamos – segundo o critério adotado – a não aceitar a hipótese nula. Portanto, a região crítica compõe-se dos valores menores que 8 e maiores que 16. Podemos ainda descrever a região com precisão maior, pois sabemos que nossa observação será um valor natural não maior que 24 (já que esta observação será o número de “caras” que aparecerão em 24 lançamentos da moeda). Logo, segue-se a seguinte descrição da região crítica:

$$\mathcal{R}_C = \{0, 1, 2, \dots, 6, 7, 17, 18, \dots, 23, 24\}$$

Resposta em (b). Introduziremos a variável aleatória que corresponde ao número de caras a serem obtidas em 24 lançamentos da moeda a serem feitos conforme “planejado” pelo enunciado do exercício. Denotaremos esta por X . Esta vai ser usada no cálculo do nível de significância do teste. Para a execução deste cálculo, precisamos agir como se fosse que a hipótese nula seja verdadeira. Isto significa que assumimos – apesar de não haver nenhuma razão real amparando tal pressuposto – que p é igual 0,5. Este pressuposto nos permite a caracterizar precisamente a distribuição de X : esta é $\text{Bin}(24; 0,5)$. A partir daí, podemos calcular a probabilidade de X

assumir seu valor dentro da região crítica do teste, o que é – pela própria definição – o nível de significância (recorde, a notação para este é α). Eis o cálculo:

$$\begin{aligned}\alpha &= \mathbb{P}[X \in \mathcal{RC}] \quad (\text{onde } X \sim \text{Bin}(24; 0, 5)) \\ &= \mathbb{P}[X \leq 7] + \mathbb{P}[X \geq 17] \\ &= \mathbb{P}[X = 0] + \cdots + \mathbb{P}[X = 7] + \\ &\quad + \mathbb{P}[X = 17] + \cdots + \mathbb{P}[X = 24]\end{aligned}$$

Para fechar a conta, usaremos os valores da distribuição binomial com $n = 24$ e $p = 0, 5$ a qual foi obtida com auxílio do pacote estatístico R. Eis a tabela de seus valores:

0	5.960464e-08	9	7.793331e-02	18	8.022547e-03
1	1.430511e-06	10	1.169000e-01	19	2.533436e-03
2	1.645088e-05	11	1.487818e-01	20	6.333590e-04
3	1.206398e-04	12	1.611803e-01	21	1.206398e-04
4	6.333590e-04	13	1.487818e-01	22	1.645088e-05
5	2.533436e-03	14	1.169000e-01	23	1.430511e-06
6	8.022547e-03	15	7.793331e-02	24	5.960464e-08
7	2.062941e-02	16	4.383749e-02		
8	4.383749e-02	17	2.062941e-02		

Somando os valores, temos que $\mathbb{P}[X \leq 7] = 0, 03195733$ e que $\mathbb{P}[X \geq 17] = 0, 01132792$. Por fim, então, temos a resposta: $\alpha = 0, 03195733 + 0, 01132792 \approx 0, 043$, ou, o nível de significância é 4,3% aproximadamente.

Resposta em (c). A pergunta (c) faz parte das perguntas do tipo “o que seria se p da moeda tivesse um valor específico”. Vale nota que fazer tal pergunta não significa saber o valor verdadeiro de p (no presente caso, por exemplo, ninguém afirma que a moeda testada possui $p = 0, 3$; é uma pergunta hipotética: “o que seria, se fosse...”).

Então, vamos fingir que $p = 0, 3$ e vamos calcular a probabilidade de que a regra de decisão erre, quer dizer, conclua que a hipótese nula é válida (sendo que a hipótese nula afirma que $p = 0, 5$, está claro onde está o erro). Vale mencionar que a conta a ser feita para $p = 0, 3$, sendo repetida para todos os valores de p diferentes de 0,5, dará a informação suficiente para deduzir a função de poder do teste. Isto revela a ligação da presente questão com o conteúdo da aula teórica onde menciona-se o Poder de Teste. Cabe, porém, a ressalva: a aula não foi além da menção e não apresentou nem o cálculo da Função de Poder de Teste, nem sua aplicação.

Nosso procedimento é muito parecido com o que usamos para o cálculo de nível de significância. Nós vamos nos colocar no eixo de tempo antes do experimento (o qual é 24 lançamentos da moeda) mas depois de ter assumido que o p da moeda é 0,3. De nosso ponto de vista, a quantidade de “caras” a aparecer neste experimento é uma variável aleatória. Denotaremos esta por Y . Temos que

$$Y \text{ tem a distribuição } \text{Bin}(24; 0, 3)$$

pois Y contará a número de “caras” em 24 lançamentos de mesma moeda, sendo que a probabilidade de obter “cara” em cada um deles é 0,3.

Com uso de Y podemos expressar a probabilidade que a regra de decisão aponte na hipótese nula:

$$\begin{aligned}\mathbb{P}[Y \text{ assumir valor fora da } \mathcal{RC}] &= \\ &= \mathbb{P}[Y = 8] + \mathbb{P}[Y = 9] + \cdots + \mathbb{P}[Y = 16] \\ &\quad \text{onde } Y \text{ tem a distribuição } \text{Bin}(24; 0, 3)\end{aligned}$$

Para fechar a conta, consultaremos a tabela da distribuição binomial com os parâmetros $n = 24$ e $p = 0, 3$. Eis esta (obtida com auxílio do pacote estatístico R):

0	1.915812e-04	9	1.221813e-01	18	6.134836e-06
1	1.970550e-03	10	7.854515e-02	19	8.302786e-07
2	9.711995e-03	11	4.284281e-02	20	8.895842e-08
3	3.052341e-02	12	1.989131e-02	21	7.261912e-09
4	6.867768e-02	13	7.869088e-03	22	4.243975e-10
5	1.177332e-01	14	2.649795e-03	23	1.581605e-11
6	1.597807e-01	15	7.570842e-04	24	2.824295e-13
7	1.760849e-01	16	1.825114e-04		
8	1.603630e-01	17	3.680902e-05		

A resposta final é 0,435 aproximadamente.

Solução do Exc. 156 a ser acrescida.

Solução do Exc. 157(a). $R_c = \{73, 74, \dots, 100\}$.

$$\alpha = \mathbb{P}[X \in R_c], \text{ onde } X \sim \text{Bin}(100; 0, 7)$$

Usando a aproximação de X pela $Y \sim \mathcal{N}(70; 21)$ temos (abaixo, $Z \sim \mathcal{N}(0, 1^2)$):

$$\alpha \approx \mathbb{P}[Y \geq 73] = \mathbb{P}\left[Z \geq \frac{73 - 70}{\sqrt{21}}\right] = 1 - 0,7422 = 0,2578$$

(b) Seja p a proporção de alunos aprovados pelo novo método e X a v.a. *número de alunos aprovados* dentro da amostra de 100 alunos monitorados. O teste será realizado para:

- $\mathbb{H}_0 : p = 0,7$
- $\mathbb{H}_A : p > 0,7$

A região crítica tem o seguinte formato:

$$\mathcal{RC} = \{c, c + 1, \dots, 99, 100\}$$

Fixando $\alpha = 10\%$, podemos determinar o valor de c a partir da relação

$$0,1 = \mathbb{P}(\text{erro I}) \tag{9.41}$$

Sabemos que, **sob a hipótese nula**, $X \sim \text{Binomial}(100; 0,7)$. Dessa forma, a equação (9.41) adquira a seguinte forma:

$$0,1 = \mathbb{P}[X \geq c], \text{ onde } X \sim \text{Binomial}(100; 0,7)$$

A v.a. X será aproximada pela v.a $Y \sim \mathcal{N}(np; np(1-p))$, onde $n = 100$ e $p = 0,7$, ou sejaq, pela $Y \sim \mathcal{N}(70; 21)$. A aproximação nos dá:

$$\begin{aligned} 0,10 &= \mathbb{P}(Y \geq c) \\ 0,10 &= \mathbb{P}\left(\frac{Y - np}{\sqrt{np(1-p)}} \geq \frac{c - 70}{\sqrt{21}}\right) \\ 0,10 &= \mathbb{P}\left(Z \geq \frac{c - 70}{\sqrt{21}}\right) \end{aligned}$$

Consultando a tabela da distribuição normal padrão teremos:

$$\begin{aligned}\frac{c - 70}{\sqrt{21}} &\approx 1,28 \\ c &\approx 1,28(\sqrt{21}) + 70 \\ c &\approx 75,87\end{aligned}$$

Portanto

$$\mathcal{RC} = \{Y : Y \geq 75,87\} = \{X : X \geq 76\}$$

Para uma proporção amostral $\hat{p} = 0,74$ devemos ter um valor $x = 74$, seque que $x \notin \mathcal{RC}$ e portanto, para um nível de significância $\alpha = 10\%$, não rejeitamos \mathbb{H}_0 , ou seja, no resultado “74 aprovados numa amostra de 100 alunos” não há evidências, ao nível de significância de 10%, de que a aprovação pelo novo método aumentou.

Solução do Exc. 158. Seja p a porcentagem dos carros que atravessam a fronteira levando contrabando após o plano ser colocado em prática. Os inspectores desejam decidir entre

$$\begin{aligned}H : p &= 0,1 \\ A : p &< 0,1\end{aligned}$$

O nível de significância é dado por

$$\mathbb{P}[X \leq 15] \text{ onde } X \sim \text{Bin}(200; 0,1)$$

Como $\mathbb{E}[X] = 200 \times 0,1 = 20$ e $\text{Var}[X] = 200 \times 0,1 \times (1 - 0,1) = 18$, então, pela aproximação de binomial por normal, temos

$$\mathbb{P}[Y \leq 15] = \mathbb{P}(Z \leq (15 - 20)/\sqrt{18}) = \mathbb{P}(Z \leq -1,17) = 1 - P(Z \geq 1,17) \approx 0,121$$

quer dizer, o nível de significância do teste é 0,121.

Solução do Exc. 159. São duas alternativas, entre as quais o dono do supermercado quer (ou precisa) escolher:

- 1-a: A marca Z de um produto está responsável por 50% das vendas do produto;
- 2-a: A marca Z está responsável por uma porcentagem maior que 50% das vendas do produto.

Se quiser, pode interpretar as alternativa da seguinte maneira:

- 1-a: A campanha promocional não mudou a popularidade da marca Z e esta continua ser responsável por 50% das vendas do produto;
- 2-a: A campanha promocional foi eficiente – conforme os promotores garantem – e a marca Z tornou-se mais popular, isto é, começou a ser responsável por uma porcentagem maior que 50% das vendas do produto.

Seja que foi a interpretação das duas alternativas, a “matematização” das mesmas é a seguinte:

(i) suponhamos que detro de cada consumidor há uma moeda que está sendo lançada cada vez que o consumidor escolha a marca do produto; (ii) suponhamos que na “cara” da moeda, o consumidor compra a marca Z, e na “coroa” compra qualquer outra marca; (iii) suponhamos que todas as moedas são independentes e têm a mesma probabilidade de dar “cara”, a qual será denotada por p . Então, a expressão matemática das duas alternativa é:

1-a: $p = 0,5$;

2-a: $p > 0,5$.

Observo que não é óbvio (para alunos) que as alternativas acima descritas em palavras tenham a expressão matemática como a de cima. Na verdade, tal expressão matemática é o nosso (os professores) modelo probabilístico para a situação na qual encontra-se o dono do supermercado. E honestamente falando, ao inventar esta situação, já partimos deste modelo probabilístico. Seja como for o caminho, alunos devem aprender a encherar exatamente este modelo por trás da descrição verbal. Pois se o modelo for outro, o método de sua análise seria diferente daquele ensinado no nosso curso.

Agora a dúvida é qual das duas alternativas levará o nome “Hipótese Nula” e qual o nome “Hipótese Alternativa”. O método da escolha neste caso é muito simples: a que tem informação exata sobre (o desconhecido) valor de p deve ser a NULA! Isto por que faremos as contas (e as conclusões baseadas nos resultados das contas), *assumindo que a hipótese nula é verdade*. Então, com este pressuposto, e com a escolha da 1-a alternativa no lugar da hipótese nula, teremos nas mãos o dado “ $p = 0,5$ ”. Com este dado, podemos fazer contas. Já no caso oposto, quer dizer, no caso da escolha da 2-a alternativa para a hipótese nula, teremos o dado “ $p > 0,5$ ”. Com este dado não dá para fazer muita conta. Resumindo, eis nossa escolha:

H: $p = 0,5$;

A: $p > 0,5$,

onde “H” significa “hipótese nula” e “A” significa “Hipótese alternativa”.

Deve ser entendido que no ambiente do Teste de Hipótese não assume-se que saberemos qual das duas é verdade de verdade, e qual é falsa de verdade. Faremos só um teste e este vai apontar para aquela das duas que ele (o teste, quer dizer) acha que seja verdadeira. Portanto, podem, a princípio, ocorrer erros de dois tipos. Estes erros estão descritos abaixo já com os nomes corretos:

Erro tipo I: É quando nosso teste aponta na hipótese alternativa como a verdadeira enquanto que na verdade a hipótese nula é verdadeira;

Erro tipo II: É quando nosso teste aponta na hipótese nula como a verdadeira enquanto que na verdade a hipótese alternativa é verdadeira.

É fácil decorrar a “numeração” dos erros: o no qual o teste erra de acertar a hipótese nula chama-se “do tipo I”, enquanto que o no qual o teste erra em acertar a hipótese alternativa chama-se “do tipo II”.

Nosso teste será baseado em número das pessoas, dentro das 18 entrevistadas, que escolham a marca Z do produto. Apriori, este número pode ser qualquer um entre 0 e 18. No enunciado está dito que se for 14 ou mais, então a decisão será que a hipótese nula é verdadeira. O conjunto dos valores que levam o teste a decidir pela veracidade da hipótese alternativa chama-se *região crítica* do teste. Portanto, $\{14, 15, 16, 17, 18\}$ é a região crítica no presente caso.

O nível de significância de teste é a probabilidade do Erro do Tipo I. Voltando à definição deste erro, podemos ver que esta é a probabilidade de que dentro das 18 entrevistadas, no mínimo 14 escolherão a marca diferente de Z, mas sendo que dentro de cada um dos 18 há moeda cuja probabilidade de dar “cara” é $p = 0,5$. (Volte acima no texto para lembrar como, dentro de nosso modelo probabilístico, pessoas escolham entre a marca Z e outras marcas do produto. Observe que adotamos $p = 0,5$ pois o Erro Tipo I significa errar na situação em qual

a hipótese nula é verdadeira. Isto implica no que todas as contas podem ser feitas com base no pressuposto que $p = 0,5$.)

Para formalizar a conta e chegar ao valor numérico, introduzimos a variável aleatória X como o número das pessoas, dentro das 18, que preferirão a marca Z, e observaremos que X tem a distribuição binomial com os parâmetros $n = 18$ e $p = 0,5$. Portanto,

$$\begin{aligned} \text{o nível de significância do teste} &= \\ \text{a probabilidade do Erro do tipo I} &= \\ &= \mathbb{P}[X \text{ pertencer à região crítica}], \text{ quando H é verdade} \\ &= \mathbb{P}[X \geq 14], \text{ quando } X \sim \text{Bin}(18; 0,5). \end{aligned}$$

Para chegar no valor, faremos aproximação de binomial pela normal. Precisamos saber que $\mathbb{E}[X] = 18 \cdot 0,5 = 9$ e $\text{Var}[X] = 18 \cdot 0,5(1 - 0,5) = 4,5 \Rightarrow \sigma[X] = 2,12$. Então:

$$\mathbb{P}[X \geq 14] = \mathbb{P}\left[Z \geq \frac{14 - 9}{2,12}\right] \mathbb{P}[Z \geq 2,39] \approx 0,01.$$

Este valor é o nível de significância do teste. Comentaria que este é muito baixo, o que é bom, pois significa que errar na escolha da hipótese nula, enquanto ela é verdadeira, tem probabilidade baixa de acontecer.

Solução do Exc. 159(c). Separei este item dos (a) e (b) pois situação aqui não pode ser misturada com a destes dois itens. Pois em (a, b) não sabemos precisamente o valor de p sob a validade da hipótese alternativa. Aqui, no (c) – sabemos sim. Como é que isto poderia ter ocorrido? Imagine que a equipe de promotores de venda fizeram um estudo meticoloso (com muito mais que as 18 pessoas mencionadas no enunciado) e descobriram que com absoluta certeza que a moeda em cada pessoa (aquela que determina a escolha entre a marca Z e as outras marcas) tornou a ter $p = 0,8$. Só que o resultado do trabalho dos promotores foi informado para o dono do supermercado por um estagiário, que confundiu-se e informou ao dono somente que p havia aumentado. Então imagine esta equipe de promotores sentados no escritório deles e pensando assim: “Ai! O dono não sabe que p foi para 0,8. Ele vai aplicar teste em 18 pessoas conforme descrito no enunciado. O teste pode dar errado (recorde: é o Erro do tipo II, em nossas nomenclatura). Caso der, o dono concluirá que nosso trabalho de promoção foi em vão. Vai ficar bravo! Com qual probabilidade isto vai ocorrer?”

Então, a pergunta acima formulada é calcular a probabilidade do erro do tipo II sob a informação adicional de que $p = 0,8$, ou, em palavras mais exatas: a probabilidade do teste decidir a favor da hipótese nula, enquanto, na verdade, $p = 0,8$. Nossas contas, portanto, serão feitas com o pressuposto de que $p = 0,8$. Aproveitamos da variável aleatória X introduzida na solução do item anterior. Só que agora tem-se: $X \sim \text{Bin}(18, 0,8)$. E outra diferença na conta em comparação com a já feita: precisamos a probabilidade de decidir a favor da hipótese nula. Recordamos que a decisão será assim caso X assumir valores não maiores que 13. Juntando tudo isto, escrevemos o que é precisa:

$$\mathbb{P}[X \leq 13], \text{ onde } X \sim \text{Bin}(18; 0,8).$$

Para achar o valor desta probabilidade, precisamos saber que $\mathbb{E}[X] = 18 \cdot 0,8 = 14,4$, $\text{Var}[X] = 18 \cdot 0,8(1 - 0,8) = 2,88 \Rightarrow \sigma[X] \approx 1,7$. Então:

$$\mathbb{P}[X \leq 13] = \mathbb{P}\left[Z \leq \frac{13 - 14,4}{1,7}\right] \mathbb{P}[Z \leq -0,82] \approx 0,2.$$

São 20% de chances de injustiçar o resultado do trabalho dos promotores em função do desconhecimento do fato que $p = 0,8$ por lado do dono de supermercado.

Observação: Eu deveria ter mais cuidado e em momentos da aproximação de binomial pelo normal, substituir “ \geq ” por “ $>$ ” e “ \leq ” por “ $<$ ”. Isto é $\mathbb{P}[X \geq 14] = \mathbb{P}[X > 15] = \mathbb{P}[Y > 15]$ e $\mathbb{P}[X \leq 13] = \mathbb{P}[X < 12] = \mathbb{P}[Y < 12]$ para apropriadas variáveis aleatórias normais Y . Refaça as contas você.

Solução do Exc. 160 a ser acrescida.

Solução do Exc. 161 a ser acrescida.

Solução do Exc. 162 a ser acrescida.

Solução do Exc. 163 a ser acrescida.

9.12.3 Soluções de exercício sobre teste de hipóteses sobre média

Solução do Exc. 164. REQUER VERIFICAÇÃO.

H: $\mu = 8$ (entrou no lugar de “ $\mu \geq 8$ ” para permitir que fizemos as contas);

A: $\mu < 8$, onde μ é a média (populacional!!!) do consumo mensal.

Faremos \bar{X}_n , a média com base em consumo mensal de n pessoas a serem escolhidas aleatoriamente. Naturalmente,

se \bar{X}_n for pequeno, é precisa aceitar A,
se \bar{X}_n for não pequeno, é precisa aceitar H.

Portanto

$$RC = [0; k].$$

O valor observado de \bar{X}_n (com $n = 80$) for de 7,4. O ND correspondente a tal observação é:

$$ND = \mathbb{P}[\bar{X}_{80} \in [0, \not\leq, \not\geq]],$$

sendo que a probabilidade calcula-se sob a veracidade da hipótese nula. Sendo que

$$\bar{X}_{80} = \frac{X_1 + \cdots + X_{80}}{80},$$

onde X_1, \dots, X_{80} são independentes (por suposição da amostragem aleatória) e $\mathbb{E}[X_i] = 8$ (por termos aceito a hipótese nula para fins do cálculo), e $\sigma[X_i] = 2,2$ (dado do enunciado), temos que

$$\bar{X}_{80} \sim \mathcal{N}\left(8; \frac{2,2}{80}\right).$$

Com isto podemos fazer contas:

$$ND = \mathbb{P}[\bar{X}_{80} \in [0, \not\leq, \not\geq]] = \mathbb{P}\left[Z \leq \frac{7,4 - 8}{0,53}\right] = \mathbb{P}[Z \leq -1,13] = 0,23.$$

ND=23% é maior que 5% desejados para o nível de significância do teste e portanto a hipótese nula deve ser aceita.

Solução do Exc. 165. REQUER VERIFICAÇÃO. O primeiro passo é identificar a distribuição cuja média está em foco de nossa atenção conforme avisamos no parágrafo acima. Bom, esta distribuição é a distribuição de horas/homem/ano perdidas em consequência de acidentes de trabalho. Você, meu querido/a aluno/a deve imaginar a seguinte situação: todas as indústrias são enumeradas de 1 a um certo número N , e em cada uma delas há o número de horas/homem/ano perdidos que segue uma certa distribuição, a mesma para todas as indústrias, porém desconhecida.

Talvez, seja demais pedir de sua imaginação que veja o número de horas/homem/ano numa indústria como uma variável aleatória. Para facilitar esta tarefa, sugiro que pense que hoje é o começo de ano e que estas horas perdidas que acontecerão no decorrer de um ano serão observados, e seu número será dividido pela quantidade de trabalhadores da indústria.

O formalismo matemático da situação, que você deve imaginar conforme meu pedido, é assim:

- (a) X_i denota a variável aleatória que corresponde ao número de horas/homem/ano perdidas na indústria i ,
- (b) as variáveis aleatórias X_1, \dots, X_n são independentes e cada uma tem a mesma distribuição.

Denotaremos por μ a média de cada variável aleatória: $\mu := \mathbb{E}[X_i]$, e denotaremos por σ^2 a variância de cada variável aleatória: $\sigma^2 := \text{Var}[X_i]$. Notarei que o exercício pergunta sobre a média, mas a variância vai ser necessária para responder a pergunta.

O exercício nos diz que a média de horas/homem/ano na década de 2000 era 60. Isto quer dizer o seguinte: em cada indústria foram observadas os valores de horas/homem/ano e foi calculada a média destes valores. Média amostral é o nome para esta. Ela era 60. Observe que a amostra era emensa pois foram analisadas todas as indústrias, e ainda mais: cada uma delas contribui com 10 valores pois a década contém 10 anos. Então naturalmente, esta média amostral deve ser vista como uma boa estimativa para a média populacional, a qual é μ , conforme nossas notações. Então, a afirmação do exercício entende-se assim: a média da distribuição de horas/homem/ano na década de 2000 era 60.

Agora, em 2011, quer dizer, foi feita uma outra amostra de horas/homem/ano, esta vez com somente 100 indústrias. (Note que as observações correspondem ao número de horas perdidos por homem no ano 2011 pois cada indústria foi observado durante um ano; é por isto que o enunciado do exercício refere-se a tais observações como “horas/homem” e não como “horas/homem/ano”.) A média desta amostra, a quantia ser denotada por $\hat{\mu}$, deu 50. Note que a amostra é relativamente pequena (cada uma das 100 indústrias contribuiu à amostra um só valor) e é por isto que a média amostral $\hat{\mu}$ pode estar diferente da média populacional μ . Isto gera a seguinte dúvida: será que a média amostral (μ , quer dizer) ficou intacta, quer dizer, seu valor continua ser 60, enquanto que o valor 50 “aconteceu” devido ‘a insuficiência do tamanho da amostra feita em 2011; ou será que a verdadeira média diminuiu (para 50, ou algo perto de 50) e esta diminuição é a razão que fez a média $\hat{\mu}$ assumir valor 50?

Em termos de teste de hipóteses, as duas alternativas formuladas acima, adquirem o seguinte formalismo:

$$\begin{aligned} H &: \mu = 60 \\ A &: \mu < 60 \end{aligned}$$

Uma vez que o exercício fornece o valor da média amostral, ficamos induzidos a testar as hipóteses com base neste valor. Daí a regra de decisão: se o valor da média amostral ficar abaixo de um certo limiar R , apontaremos em A como a verdade, enquanto que se este valor for acima de R , apontaremos em H como verdade.

Com isto, o erro do tipo I torna a ser: aceitar a hipótese A enquanto que na verdade a hipótese H é verdade. A aceitação de A acontece quando a média amostral de 100 indústrias fica abaixo de R :

$$\frac{X_1 + X_2 + \cdots + X_{100}}{100} \leq R,$$

enquanto que a verossimilhança de H significa que cada X tem a média (isto é, $\mathbb{E}[X_i]$) igual a 60. Acontece que a informação sobre a média de cada X não é suficiente para fazer os cálculos. Precisaremos do seguinte fato:

A variável aleatória

$$\frac{X_1 + X_2 + \cdots + X_{100}}{100}$$

tem distribuição aproximadamente normal com a média μ e a variância $\sigma^2/100$, onde μ e σ^2 são, respectivamente, a média e a variância de cada X . No nosso caso tal média é 50 (por ter “aceitado” a hipótese H), e a variância é igual a 20^2 (conforme os dados do enunciado do exercício).

Com isto podemos agora dar expressão matemática ao nível de significância do teste:

$$\begin{aligned} \text{o nível de significância} &= \mathbb{P} \left\{ \frac{X_1 + X_2 + \cdots + X_{100}}{100} \leq R \right\} \\ &= \mathbb{P} \{ Y \leq R \}, \end{aligned}$$

onde Y é a variável aleatória Normal com a distribuição $\mathcal{N}\left(60; \frac{20^2}{100}\right)$.

Pelo anunciado do exercício (sua pergunta (c)), precisamos achar R correspondente ao nível de significância 10%. Isto significa que

$$0,1 = \mathbb{P}\{Y \leq R\} = \mathbb{P}\left\{\frac{Y - 60}{20/\sqrt{100}} \leq \frac{R - 60}{20/\sqrt{100}}\right\} = \mathbb{P}\left\{Z \leq \frac{R - 60}{20/\sqrt{100}}\right\},$$

onde Z é a variável aleatória normal padrão.

Da tabela da distribuição normal padrão, temos que

$$\mathbb{P}\{Z \leq 1,29\} \approx 0,9,$$

o que implica no que

$$0,1 = \mathbb{P}\{Z \leq -1,29\},$$

e portanto,

$$\frac{R - 60}{20/\sqrt{100}} = -1,29,$$

de onde

$$R = 57,42$$

Vamos agora à observação, quer dizer, vamos lembrar que $\hat{\mu} = 50$. Sendo que $50 < 57,42$, devemos rejeitar a hipótese nula, quer dizer, a amostragem feita em 2011 indica, com o nível de significância de 10%, a diminuição da média da distribuição de horas/homem/ano no 2011 em relação à média na década de 2010.

Solução do Exc. 166. A acrescentar.

Solução do Exc. 167. REQUER VERIFICAÇÃO.

Temos que X representa o tempo de vida de galinhas de uma fazenda, em anos. Queremos testar as seguintes hipóteses para X :

- \mathbb{H}_0 : O tempo de vida da população tem distribuição $\mathcal{N}(10; 1^2)$
- \mathbb{H}_A : O tempo de vida da população tem distribuição $\mathcal{N}(\mu; 1^2)$, para $\mu > 10$

Queremos, portanto, testar se o tempo médio de vida das galinhas, μ , seja igual aos 10 anos ou maior que 10 anos.

Para realizar o teste, obtemos uma amostra de 400 galinhas. Com um nível de significância $\alpha = 0,08$ e, levando-se em conta a hipótese alternativa, teremos uma região crítica para o teste da seguinte forma:

$$\mathcal{RC} = \{X : \bar{X} > c\}$$

Fixando $\alpha = 8\%$, calculemos o valor c :

$$\begin{aligned} \alpha &= \mathbb{P}(\text{erro I}) \\ 0,08 &= \mathbb{P}(\text{rejeitar } \mathbb{H}_0 | \mathbb{H}_0 \text{ é verdadeira}) \\ 0,08 &= \mathbb{P}(\bar{X} \in \mathcal{RC} | \mu = 10) \\ 0,08 &= \mathbb{P}(\bar{X} > c | \mu = 10) \end{aligned} \tag{9.42}$$

Sabemos que, **sob a hipótese nula**, $\bar{X} \sim \mathbb{N}(10; \frac{1}{400})$. Assim:

$$\begin{aligned} 0,08 &= \mathbb{P}(\bar{X} > c) \\ 0,08 &= \mathbb{P}\left(\frac{\bar{X} - \mu}{\sigma} > \frac{c - 10}{\sqrt{\frac{1}{400}}}\right) \\ 0,08 &= \mathbb{P}\left(Z > \frac{c - 10}{\sqrt{\frac{1}{400}}}\right) \end{aligned}$$

Consultando a tabela da distribuição normal padrão teremos:

$$\begin{aligned} \frac{c - 10}{\sqrt{\frac{1}{400}}} &\approx 1,41 \\ c &\approx \frac{1,41}{20} + 10 \\ c &\approx 10,1 \end{aligned}$$

Portanto

$$\mathcal{RC} = \{\bar{X} : \bar{X} > 10,1\}$$

Observamos um tempo de vida média amostral de 10,5 anos e, dessa forma, o valor observado encontra-se na região crítica do teste. Portanto, há indícios, a um nível de significância de 8%, para se rejeitar a hipótese nula. Isto é, concluímos que o tempo médio de vida das galinhas aumentou após começarem a receber o alimento suplementar.

Solução do Exc. 168. REQUER VERIFICAÇÃO. Temos que X representa o peso dos pacotes de café produzidos pela máquina, em Kg. Queremos averiguar se a máquina está desregulada, ou seja, se o peso médio dos pacotes de café diminuiu. Realizaremos o seguinte teste de hipóteses:

- \mathbb{H}_0 : O peso dos pacotes produzidos pela máquina tem distribuição $\mathcal{N}(0, 5; 0,05^2)$
- \mathbb{H}_A : O peso dos pacotes produzidos pela máquina tem distribuição $\mathcal{N}(\mu; 0,05^2)$,

Temos uma amostra com 100 pacotes e obteremos a região crítica do teste a um nível de significância de 3%. Com base na hipótese alternativa, a região crítica do teste é da seguinte forma:

$$\mathcal{RC} = \{X : \bar{X} < c\}$$

Fixando $\alpha = 3\%$, calculemos o valor c ::

$$\begin{aligned} \alpha &= \mathbb{P}(\text{erro I}) \\ 0,03 &= \mathbb{P}(\text{rejeitar } \mathbb{H}_0 | \mathbb{H}_0 \text{ é verdadeira}) \\ 0,03 &= \mathbb{P}(\bar{X} \in \mathcal{RC} | \mu = 10) \\ 0,03 &= \mathbb{P}(\bar{X} < c | \mu = 10) \end{aligned} \tag{9.43}$$

Sabemos que, **sob a hipótese nula**, $\bar{X} \sim \mathbb{N}(0, 5; \frac{0,05^2}{100})$. Assim:

$$\begin{aligned} 0,03 &= \mathbb{P}(\bar{X} > c) \\ 0,03 &= \mathbb{P}\left(\frac{\bar{X} - \mu}{\sigma} > \frac{c - 10}{\sqrt{\frac{1}{400}}}\right) \\ 0,03 &= \mathbb{P}\left(Z > \frac{c - 0,5}{\sqrt{\frac{0,05^2}{100}}}\right) \end{aligned}$$

Consultando a tabela da distribuição normal padrão teremos:

$$\frac{c - 0,5}{\sqrt{\frac{0,05^2}{100}}} \approx -1,88$$

$$c \approx \frac{-1,88}{10} + 0,5$$

$$c \approx 0,491$$

Portanto

$$\mathcal{RC} = \{\bar{X} : \bar{X} < 0,491\}$$

Se observarmos um peso médio amostral dos pacotes igual ao 0,498 kg, concluímos que não há indícios, a um nível de significância de 3%, para rejeitarmos a hipótese nula. Portanto, concluímos que o peso médio dos pacotes de café produzidos pela máquina não diminuiu, isto é, permaneceu igual à 0,5 kg.

Solução do Exc. 169. A ser colocada.

Solução do Exc. 170. REQUER VERIFICAÇÃO. Este é um bom exercício que merece uma explicação adicional voltada ao esclarecimento do surgimento da situação discutida nele. Imagine então que há um graaande lote de lâmpadas e que o fabricantes destas afirma que o tempo da duração média da vida das lâmpadas do lote é de 400 horas. Com tal afirmação o fabricante quer dizer o seguinte: se pegarmos todas as lâmpadas do lote, testarmos cada uma delas e anotarmos o tempo de sua vida (até se queimar), e se fizermos a média das “vidas” das lâmpadas testadas, então obtiveremos o valor 400. Em palavras científicas, 400 é a média populacional da vida. Pois bem, aconteceu que por alguma razão escolhemos aleatoriamente 25 lâmpadas, observamos os tempos da vida delas e fizemos a média das observações; deu 395. Em palavras científicas, 395 é a média amostral da vida. Eu usarei a notação \bar{x}_{25} para este número, assinalando com esta que á o valor médio (a barra acima de x significa “média”, em jargão de estatísticos) e tal valor foi obtido com base numa amostra de tamanho 25 (este afo está assinalado pelo indice 25 do x).

Será que 395 da média amostral dismente a afirmação de que a média populacional é 400? Não é possível confirmar isto sem análise adequada, pois reparamos aquí com uma situação parecida com a que vimos no caso de lançamentos de moeda honesta: a “honestidade” da moeda significa que o número médio de caras obtidas numa série longa de lançamentos deve ser 50%. Mas se lancarmos a moeda 20 vezes, o número de caras a serem vistas nestes lançamentos pode muito bem ser diferente de 10. Por exemplo, ao observar 8 caras e 12 coroas, não teríamos muita desconfiança da honestidade da moeda pois atribuiríamos tal separação desigual às flutuações aleatórias. No caso da lâmpadas, a situação é muito similar: a média amostral ficou abaixo da média populacional anunciada pelo fabricante. Mas será que esta diferença permita a alegar que o anuncio do fabricante é falso?

O conteúdo da parágrafo anterior ampara a seguinte vizão na situação: temos que escolher entre a hipótese

H: a vida média populacional é 400, ou em notação científica, $\mu = 400$,

e a hipótese

A: a vida média populacional é menor que 400, ou em notação científica, $\mu < 400$.

Observe que introduzi a notação μ para a vida média populacional, e que já atribui o nome “nula” para a primeira das duas hipótese e o nome “alternativa” para a segunda delas. Observe também, que a hipótese alternativa alega “ $\mu < 400$ ” e não “ $\mu > 400$ ” e a razão para tal é que

a média amostral acusou valor menor que 400. Vale notar, que a princípio, poderíamos testar a hipótese “H: $\mu = 400$ ” contra a hipótese “A: $\mu > 400$ ”, mas se fizéssemos isto, então com a valor observado 395 para a média amostral, o teste de hipóteses confirmaria a hipótese H com qualquer nível de significância menor que 50%. Você pode refazer minhas contas para confirmar esta afirmação, mas isto levaria você para fora do conjunto mínimo de conhecimentos necessários para a aprendizagem satisfatória da disciplina MAE116.

Vamos à construção do teste de hipóteses, quer dizer, o procedimento a ser usado para escolher uma das duas hipóteses formuladas acima. Nota-se de passagem que nada garante que tal procedimento escolhe obrigatoriamente aquela das duas que é verdadeira. A escolha deve se basear no valor da média amostral de vida, pois este é o único dado fornecido no enunciado. Ainda mais: tal dado refere-se à média calculada na amostra de tamanho 25. Portanto, ao construir nosso teste teremos em mente que ele usará o valor da média de amostra de 25 lâmpadas. Chamaremos este por \bar{X}_{25} . Observe o uso da letra maiúscula, pois em nossos argumentos, até um certo ponto, este valor será visto como uma variável aleatória. E o argumento é assim: Há discordância entre as hipóteses sobre o valor da média populacional. Uma alega que esta é menor que 400, outra - que é 400. Tal discordância e nossa esperança de que a média amostral seja uma boa aproximação para a média populacional sugerem em conjunto que o teste deve ser assim:

caso o valor de \bar{X}_{25} for *pequeno* então concluiremos que $\frac{A}{H}$ é verdadeira.
grande

Para separar “pequenos” dos “grandes”, introduziremos o limiar k que entra em nossa regra da seguinte maneira:

caso o valor de \bar{X}_{25} for $\leq k$ então concluiremos que $\frac{A}{H}$ é verdadeira.
 $> k$

O valor numérico de k depende do nível de significância desejado. O vínculo entre os dois dá-se pela seguinte conta, na qual o nível de significância foi tomado de 0,05 conforme solicitado no item (b) do exercício.

$$\begin{aligned} 0,05 &= \text{a probabilidade do erro do tipo I} \\ &= \text{a probabilidade de aceitar A enquanto H é verdadeira} \\ &= \mathbb{P}[\bar{X}_{25} \leq k], \text{ quando } \mu = 400 \end{aligned}$$

Nossa atenção agora é com o cálculo da $\mathbb{P}[\bar{X}_{25} \leq k]$ sob a condição de que $\mu = 400$. Recordamos que \bar{X}_{25} será calculado como a média com base em 25 observações de vida de 25 lâmpadas. Portanto, ao usar a notação X_i para a vida da i -ésima lâmpada, podemos escrever a relação

$$\bar{X}_{25} = \frac{X_1 + \cdots + X_{25}}{25}.$$

O ingrediente importante para nossas contas é que $\mathbb{E}[X_i] = 400$ para todo $i = 1, 2, \dots, 25$. Isto é a consequência direta da aceitação da hipótese nula em nossa presente conta. (Note que não sabemos e não poderemos saber com certeza absoluta se a hipótese nula é verdadeira. Então, quanto à veracidade da hipótese nula, nos assumimos esta sem nenhum teste. Tal suposição é necessária para a conta do nível de significância.) Observe que há interpretação dupla da hipótese nula. Numa das versões desta interpretação, a que foi no começo da solução, dizemos que a média populacional de vida das lâmpadas é 400. Na outra, a ser usada agora, dizemos que 400 será a esperança matemática da variável aleatória que representa a vida de uma lâmpada.

O segundo ingrediente importante para o sucesso de nossa conta está dado no enunciado: “o desvio padrão de tempo de vida das lâmpadas é igual a 30 horas”. Isto significa que $\sigma[X_i] = 30$ para todo $i = 1, 2, \dots, 25$.

O conhecimento das médias e das variâncias de cada uma das variáveis aleatórias X_1, \dots, X_{25} e a fórmula da transparência 6 da aula 9 nos dá que

\bar{X}_{25} tem distribuição (aproximadamente) normal com média 400 e variância $\frac{(30)^2}{25}$.

Então, com Z denotando a variável aleatória Normal Padrão, temos que

$$0,05 = \mathbb{P}[\bar{X}_{25} \leq k] = \mathbb{P}\left[Z \leq \frac{k - 400}{30/5}\right],$$

o que implica, via a tabela da distribuição Normal padrão que $\frac{k-400}{30/5} = -1,65$, de onde $k = 390,1$.

Então o valor do limiar k foi estabelecido. Ele tem vínculo direto com o conceito de “Região Crítica”. Pela sua definição, esta é a região que contempla os valores de \bar{X}_{25} para os quais será aceita a hipótese alternativa. Então, a região crítica é $(-\infty; 390,1]$. Esta resposta pode ser melhorada um pouco: devido ao fato que \bar{X}_{25} pode assumir só valores positivos, podemos então refinar a resposta: a Região Crítica no caso é $(0; 390,1]$.

E agora, qual das duas hipóteses vence sendo julgadas pela regra de decisão contruída? É simples, a resposta. O valor observado de \bar{X}_{25} é 395. Este fica acima do limiar da regra $k = 390,1$ (ou em outras palavras, fica fora da Região Crítica do teste), e portanto a hipótese nula deve ser aceita como a verdadeira.

(c) Aquí, sem nenhuma justificativa, está sendo alegado que a verdade é que $\mu = 385$. Pergunta-se então o valor da probabilidade do erro do tipo II. São duas situações compatíveis com esta realidade. 1) Que nos sabemos a verdade, mas quem aplica o teste não sabe dela e daí surge a curiosidade nossa em achar a probabilidade de que ao aplicar o teste a pessoa erra. 2) Que não sabemos o verdadeiro valor de μ mas queremos saber a probabilidade de cometer erro caso este fosse 385. Em qualquer um dos dois casos, a conta é a mesma.

A probabilidade do erro do tipo II é a probabilidade de decidir que a hipótese nula é verdade enquanto que na verdade a hipótese alternativa é. Lembrando e expressão da regra de decisão em termos dos valores da variável aleatória \bar{X}_{25} , temos a seguinte expressão da probabilidade do erro

$$\mathbb{P}[\bar{X}_{25} > 390,1] \text{ quando a hipótese alternativa é verdade, isto é, quando } \mu < 400.$$

Acontece que assumimos que $\mu = 385$. Por isto, “ $\mu < 400$ ” na expressão acima deve ser substituído por “ $\mu = 385$ ”. Lembramos que μ é a esperança de cada um das X_i 's que compõem \bar{X}_{25} , ou, em termos matemáticos, $\mathbb{E}[X_i] = 385$ para todo $i = 1, 2, \dots, 25$. Lembramos ainda que $\sigma[X_i] = 30$ para todo $i = 1, 2, \dots, 25$. Isto faz com que no presente caso

\bar{X}_{25} tem distribuição (aproximadamente) normal com média 385 e variância $\frac{(30)^2}{25}$.

Daí, com Z sendo a notação para a variável aleatória Normal Padrão, temos que

$$\mathbb{P}[\bar{X}_{25} > 390,1] = \mathbb{P}\left[Z > \frac{390,1 - 385}{30/5}\right] = \mathbb{P}[Z > 0,85] = 1 - 0,8023 \approx 0,2.$$

Solução do Exc. 171.