Departamento de Física Experimental

Métodos Estatísticos de Física Experimental

Probabilidade

(*Exoesqueleto de Cigarra*)

25-26 de fevereiro de 2014

Paulo R. Pascholati

Exoesqueleto de Cigarra - Ibiuna

Sumário

Prólogo

Aqui são apresentados alguns parâmetros de interesse na análise de resultados experimentais e também conceitos de probabilidade utilizados na disciplina.

Em zoologia, chama-se exoesqueleto à cutícula resistente, mas flexível, que cobre o corpo de muitos animais e protistas, fornecendo proteção para os órgãos internos, suporte para os músculos e evita também a perda de água.

Cigarra, classificação científica ou taxinomia (sistema de Karl von Linnée (ou

Sumário

Distribuição

Distribuição da População e Distribuição Amostral

A distribuição da população seria obtida se fosse possível realizar infinitas medições da quantidade x. A distribuição amostral é obtida realizando um número finito de medições da quantidade x. Os parâmetros da distribuição da população serão denotados por letras do alfabeto grego e por letras do alfabeto romano aqueles estimados por meio da distribuição amostral.
Obtem-se os parâmetros da distribuição da população supondo que os resultados experimentais se aproximam dela quando o número de medições se aproximam de infinito, assim;

$$(parametro\ da\ populacao) = \lim_{x \to +\infty} (parametro\ experimental)$$

Distribuição da População e Distribuição Amostral
Média

A média $\overline{x}$ de uma distribuição experimental de N observações da quantidade x é:

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

e a media μ da distribuição da população é definida como:

$$\mu = \lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)$$

Distribuição da População e Distribuição Amostral

Mediana

A mediana da distribuição da população $\mu_{1/2}$ é definida como

$$P(x_i < \mu_{1/2}) = P(x_i > \mu_{1/2})$$

A linha da mediana divide a área da distribuição de probabilidade em metades.

A mediana para o caso de N resultados é calculada como

- para N ímpar $mediana = Y_{((N+1)/2)}$; e
- para N par $mediana = (Y_{N/2} + Y_{(N/2)+1})/2$.

Distribuição da População e Distribuição Amostral
Moda

A moda ou o valor mais provável x_{max} é aquele em que a distribuição da população tem o maior valor é é definida como

$$P(\mu_{max}) \geq P(x \neq \mu_{max})$$

A linha da mediana divide a área da distribuição de probabilidade em metades.

Distribuição da População e Distribuição Amostral

Moda, Média e Mediana na Função LogNormal

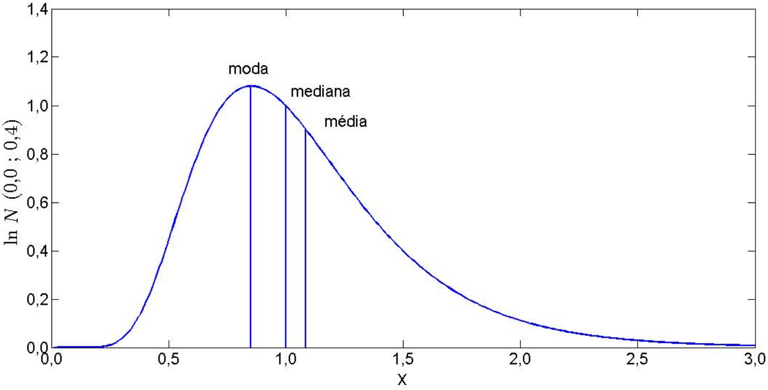

Figura 1. Função log-normal com $\mu = 0,0$ e $\sigma = 0,4$.
Média = 1,08; mediana = 1,00 e moda = 0,82.

Distribuição da População e Distribuição Amostral

Moda, Média e Mediana na Função LogNormal

$$lnN(x|\mu,\sigma) = \frac{1}{x\sqrt{2\pi\sigma}} e^{\frac{(ln\ \ x-\mu)^2}{2\sigma^2}}$$

com $\sigma > 0$, $\mu \in \Re$ em escala logarítmica e $x \in (0, +\infty)$.

$$média = e^{\mu+\sigma^2/2}$$

$$mediana = e^{\mu}$$

$$moda = e^{\mu-\sigma^2}$$

Sumário

1. Prólogo
2. Distribuição
   - Distribuição da População e Distribuição Amostral
   - Média
   - Mediana
   - Moda
   - Função LogNormal
3. Probabilidade
   - Variável Discreta
   - Variável Contínua

Probabilidade
Variável Discreta

Considere um processo cujos m possíveis eventos são: $y_1, y_2 \cdots y_m$ e os resultado obtidos em N repetições do processo $Y_1, Y_2 \cdots Y_n$ e $N(y_i)$ a frequência de ocorrência do evento y_i tem-se

$$N = \sum_{i=1}^{m} N(y_i)$$

A frequência relativa de ocorrer o evento y_i é

$$F(y_i) = \frac{N(y_i)}{N}$$

e a probabilidade de ocorrência desse mesmo evento y_i é

$$P(y_i) = \lim_{N \to \infty} F(y_i) = \lim_{N \to \infty} \frac{N(y_i)}{N}$$

Probabilidade

Variável Discreta – Propriedade

A soma da frequência de todos eventos possíveis é igual a unidade.

$$\sum_{i=1}^{m} F(y_i) = \sum_{i=1}^{m} \frac{N(y_i)}{N}$$

$$\frac{1}{N} \sum_{i=1}^{m} N(y_i) = \frac{N}{N} = 1$$

idem da probababilidade

$$\sum_{i=1}^{m} P(y_i) = 1$$

Probabilidade

Variável Discreta – Média

A média de N medições ($j = 1 \to N$) pode ser obtida por

$$\overline{y} = \frac{1}{N}\sum_{j=1}^{N} Y_j = \frac{1}{N}\sum_{i=1}^{m} \frac{N(y_i)}{N} \cdot y_i$$

$$= \sum_{i=1}^{m} \frac{N(y_i)}{N} \cdot y_i = \sum_{i=1}^{m} F(y_i) \cdot y_i$$

Probabilidade

Variável Discreta – Valor verdadeiro

$$\mu = \lim_{N\to\infty} \sum_{i=1}^{m} F(y_i) \cdot y_i$$

$$= \sum_{i=1}^{m} \lim_{N\to\infty} F(y_i) \cdot y_i = \sum_{i=1}^{m} P(y_i) \cdot y_i$$

O valor verdadeiro, μ, do mensurando y

$$\mu = \lim_{N\to\infty} \overline{y}$$

A média $\overline{y}$ é um estimador do valor verdadeiro μ.

Probabilidade

Variável Discreta – Variança e Desvio Padrão

A variança da população, também conhecida como segundo momento, é escrito como

$$\sigma^2 = \lim_{N\to\infty} \sum_{i=1}^{m} F(y_i) \cdot (y_i - \mu)^2$$

$$= \lim_{N\to\infty} \sum_{i=1}^{m} F(y_i) \cdot (y_i^2 - 2y_i\mu + \mu^2)$$

$$= \lim_{N\to\infty} \left(\sum_{i=1}^{m} F(y_i) \cdot y_i^2 - \sum_{i=1}^{m} F(y_i) \cdot 2y_i\mu + \sum_{i=1}^{m} F(y_i) \cdot \mu^2 \right)$$

$$= \lim_{N\to\infty} \left(\sum_{i=1}^{m} F(y_i) \cdot y_i^2 \right) - 2\mu^2 + \mu^2 = \lim_{N\to\infty} \left(\sum_{i=1}^{m} F(y_i) \cdot y_i^2 \right) - \mu^2$$

Probabilidade

Variável Discreta – Variança e Desvio Padrão

O desvio padrão da população é $\sigma = \sqrt{\sigma^2}$.
A variança da amostra, também segundo momento, é escrito como[1]

$$s^2 = \frac{1}{N-1}\sum_{j=1}^{N}(Y_j - \overline{y})^2$$

$$= \frac{N}{N-1}\sum_{i=1}^{m} F(y_i)\cdot(y_i - \overline{y})^2$$

O fator $N-1$ aparece ao invés de N porque o valor de $\overline{y}$ é obtido dos dados. O desvio padrão da amostra é $s = \sqrt{s^2}$.

[1] Lembre-se de que y_i são os eventos possíveis de ocorrer no processo e Y_j é o resultado da medição j.

Probabilidade
Variável Contínua

No caso em que os possíveis eventos são valores de variável contínua no intervalo $[Y_{min}, Y_{máx}]$, é necessário definir uma **função densidade de probabilidade – fdp** para o processo:

$$p(Y) = \lim_{\Delta Y \to 0} \frac{P(Y)}{\Delta Y}$$

O valor verdadeiro é

$$\mu = \int_{Y_{min}}^{Y_{máx}} Y \; p(Y) \; dY$$

e a variança

$$\sigma^2 = \int_{Y_{min}}^{+\infty} p(Y)(Y - \mu)^2 dY = \int_{Y_{min}}^{Y_{máx}} p(Y) \; Y^2 \; dY - \mu^2$$