LCE0216
Introdução à Bioestatística Florestal

5. Principais Modelos Discretos

Profa. Dra. Clarice Garcia Borges Demétrio

Monitor: Silvio Gomes

Escola Superior de Agricultura "Luiz de Queiroz"
Universidade de São Paulo

Piracicaba, 30 de março de 2020

Modelos discretos

- Estabelecem a relação entre variável e a realização do experimento que a origina;
- Uma variável aleatória segue determinado modelo se cada possível valor da variável acontece conforme uma determinada lei de atribuição de probabilidades;
- A lei de atribuição é dada pela **função de probabilidade**;
- Para alguns casos a função de probabilidade pode ser escrita de maneira mais compacta. Esses casos refletem variáveis aleatórias que ocorrem com frequência em situações práticas.
- Neste curso, veremos os modelos discretos **Bernoulli**, **binomial** e **Poisson**.

Distribuição de Bernoulli

Definição

Experimentos que admitem apenas dois resultados possíveis (sucesso ou fracasso) recebem o nome de ensaios de Bernoulli e originam uma variável aleatória com distribuição de Bernoulli.

Experimento:

O departamento de Entomologia e Acarologia da ESALQ/USP realizou um experimento para verificar a eficácia de um novo produto para controle de determinada praga. Um grupo de 30 insetos foram submetidos à nova substância e, depois de um determinado período, foram avaliados. Tomando-se ao acaso, um inseto do estudo, verifica-se se este está vivo ou morto.

Distribuição de Bernoulli

Variável aleatória X: mortalidade da praga.

$$X = \begin{cases} x = 1, & \text{se morreu} \\ x = 0, & \text{se não morreu} \end{cases}$$

Algumas pressuposições:

- É realizada apenas uma repetição do experimento;
- Apenas dois resultados possíveis: morreu ou não morreu.

Distribuição de Bernoulli

Evento M = {O inseto morreu}

$$P(M) = \pi \qquad P(\bar{M}) = 1 - \pi.$$

Distribuição de probabilidade:

Resultados	x	$P(X = x)$
$\bar{M}$	0	$1 - \pi$
M	1	π
Total		$(1 - \pi) + \pi = 1$

Portanto, a variável aleatória X: mortalidade, tem distribuição de Bernoulli.

Distribuição de Bernoulli

Função de probabilidade

Seja X uma variável aleatória com distribuição de Bernoulli com probabilidade de sucesso π, em que $X = 1$ se o resultado é sucesso e $X = 0$ se o resultado é fracasso. Então, a função de probabilidade de X fica dada por

$$P(X = x) = \pi^x (1 - \pi)^{1-x}.$$

Logo,

$$P(X = 0) =$$

$$P(X = 1) =$$

Denota-se por $X \sim \text{Be}(\pi)$.

Distribuição de Bernoulli

Função de probabilidade

Seja X uma variável aleatória com distribuição de Bernoulli com probabilidade de sucesso π, em que$X = 1$ se o resultado é sucesso e $X = 0$ se o resultado é fracasso. Então, a função de probabilidade de X fica dada por

$$P(X = x) = \pi^x (1 - \pi)^{1-x}.$$

Logo,

$$P(X = 0) = \pi^0 (1 - \pi)^{1-0} = 1 - \pi.$$

$$P(X = 1) = \pi^1 (1 - \pi)^{1-1} = \pi.$$

Denota-se por $X \sim \text{Be}(\pi)$.

Distribuição de Bernoulli

Média

A esperança (ou valor médio) da distribuição de Bernoulli é dada por

$$\mathrm{E}(X) = \mu_X = \sum_{x=0}^{1} xP(X = x) = 0 \times (1 - \pi) + 1 \times \pi = \pi.$$

Variância

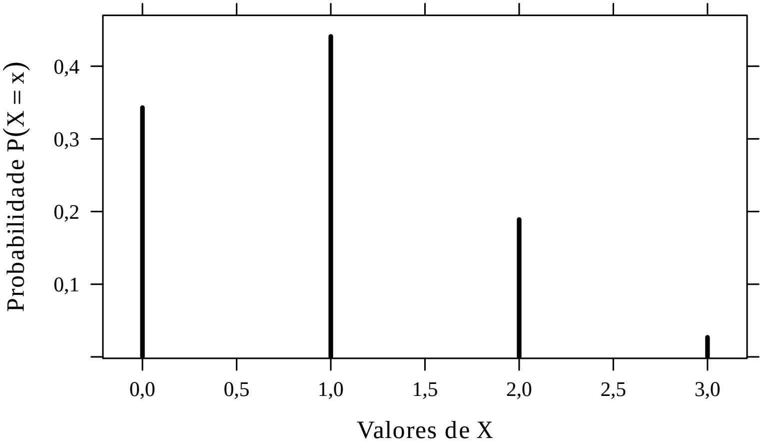
A variância de X é definida por

$$\mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}(X^2) - \left[\mathrm{E}(X)\right]^2 = \pi - \pi^2 = \pi(1 - \pi)$$

Logo, o desvio padrão de uma variável aleatória com distribuição de Bernoulli é dado por:

$$\sigma_X = \sqrt{\pi(1 - \pi)}.$$

Representação gráfica



Exemplo: Um pesquisador diz que o tratamento das estacas com uma certa concentração de hormônio eleva a porcentagem esperada de enraizamento. 10 estacas foram tratadas e destas, 6 enraizaram. Escolhe-se ao acaso uma estaca. Seja X = "a estaca enraizar", verifique se é um ensaio de Bernoulli. Determinar a $P(X = x)$, calcular $\mathrm{E}(X)$ e $\mathrm{Var}(X)$.

Solução

$\pi = 0,6$

$P(X = x) = (0,6)^x (0,4)^{(1-x)}$

$\mathrm{E}(X) = \pi = 0,6$

$\mathrm{Var}(X) = \pi(1 - \pi) = 0,6 \times 0,4 = 0,24$

Distribuição binomial

Motivação
Experimento: Verificar se dois insetos submetidos à uma nova substância permaneceram vivos ou morreram.
Pressuposições:

- O fato de um inseto morrer, ou não, não tem influência no fato de o outro inseto morrer, ou não; ou seja, as mortes são **independentes**;
- A probabilidade de os insetos morrerem é a mesma, igual a π.
- Só há dois resultados possíveis para cada inseto: morreu ou não morreu (ensaio de Bernoulli); e
- Existem duas repetições.

Distribuição binomial

Variável aleatória X = número de insetos mortos.

Resultado	Probabilidade	x
MM	$\pi\pi$	2
$M\bar{M}$	$\pi(1-\pi)$	1
$\bar{M}M$	$(1-\pi)\pi$	1
$\bar{M}\bar{M}$	$(1-\pi)(1-\pi)$	0
Total	1	

Distribuição de probabilidades

x_i	$P(X=x_i)$
0	$(1-\pi)^2$
1	$2\pi(1-\pi)$
2	π^2
Total	1

Distribuição binomial

Generalizando...

A probabilidade de x insetos morrerem e, portanto, $n - x$ insetos permanecerem vivos, nesta sequência,

$$\underbrace{MM\ldots M}_{x},\underbrace{\bar{M}\bar{M}\ldots\bar{M}}_{n-x}$$

é dada por $\pi^x(1-\pi)^{n-x}$.

Mas note que outras sequências podem ocorrer com a mesma probabilidade, tais como:

$$MMM\ldots\bar{M}\bar{M}M\bar{M}\ldots\bar{M} \quad \text{ou} \quad MMM\ldots\bar{M}M\bar{M}\bar{M}\ldots\bar{M}.$$

Existem

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

de tais sequências.

Distribuição binomial

Generalizando...

Logo,

$$P(X = x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x}, \quad \text{para } x = 0, 1, 2, \ldots, n.$$

Observações:

- A denominação binomial decorre do fato de os coeficientes $\binom{n}{x}$ serem exatamente os coeficientes do desenvolvimento binomial dos termos $(a + b)^n$;
- O cálculo dos coeficientes, para n e x grandes, é difícil de ser realizado.

Notação: $X \sim B(n; \pi)$.

Distribuição binomial

Pressuposições:

- Existem n repetições ou provas idênticas do experimento;
- Só há dois tipos de resultados possíveis em cada repetição;
- As probabilidades π de sucesso e $(1 - \pi)$ de fracasso permanecem constantes em todas as repetições;
- Os resultados das repetições são independentes uns dos outros.

Nota: No caso de alelopatia, isso não ocorre e a distribuição binomial não é adequada.

Distribuição binomial

Distribuição binomial

A variável aleatória X correspondente ao número de sucessos em n ensaios de Bernoulli independentes (no sentido probabilístico) e com mesma probabilidade π de sucesso em cada ensaio, tem distribuição binomial com parâmetros n e π. A função de probabilidade de X é expressa na forma

$$P(X = x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}, \quad \text{para } x = 0, 1, 2, \ldots, n.$$

Notação: Denota-se $X \sim B(n, \pi)$

Distribuição binomial

Média

Se $X \sim B(n, \pi)$ pode-se escrever $X = X_1 + X_2 + \ldots + X_n$, em que $X_i \sim Be(\pi)$, para $i = 1, \ldots, n$, independente e identicamente distribuídas. Assim, obtém-se

$$\mu_X = \mathrm{E}(X) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + \ldots + \mathrm{E}(X_n) = n\pi.$$

Variância

De forma semelhante, como se têm n ensaios independentes, então

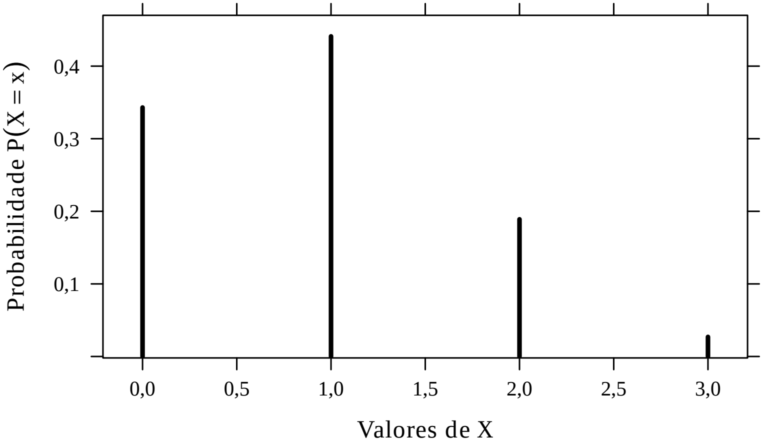
$$\sigma_X^2 = \mathrm{Var}(X) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n) = n\pi(1-\pi)$$

Desvio-padrão

Daí segue que

$$\sigma_X = \sqrt{\mathrm{Var}(X)} = \sqrt{n\pi(1-\pi)}$$

Representação gráfica



Distribuição binomial

Exemplo: Um lote de *Eucaliptus saligna* com uma proporção de 5% de sementes híbridas (*E. saligna* $\times$ *E. cloeziana*) foi utilizado para a implantação de uma floresta. Se dez árvores desta floresta forem selecionadas ao acaso, qual a probabilidade de

(a) Nenhuma delas ser híbrida;

(b) Pelo menos uma delas ser híbrida;

(c) Todas elas serem híbridas.

(d) Seja X o número de sementes híbridas em 10 sementes. Calcule $\mathrm{E}(X)$ e $\mathrm{Var}(X)$.

Exemplo: Um lote de *Eucaliptus saligna* com uma proporção de 5% de sementes híbridas (*E. saligna* $\times$ *E. cloeziana*) foi utilizado para a implantação de uma floresta. Se dez árvores desta floresta forem selecionadas ao acaso, qual a probabilidade de

(a) Nenhuma delas ser híbrida;

(b) Pelo menos uma delas ser híbrida;

(c) Todas elas serem híbridas.

(d) Seja X o número de sementes híbridas em 10 sementes. Calcule $\mathrm{E}(X)$ e $\mathrm{Var}(X)$.

Solução

$\pi = 0,05 \quad n = 10$

(a) $P(X = 0) = \binom{n}{x}\pi^x(1-\pi)^{n-x} = \binom{10}{0}(0,05)^0(0,95)^{(10-0)} = 0,5987369$

(b) $P(X \geq 1) = P(X = 1) + P(X = 2) + \ldots + P(X = 10) = 1 - P(X = 0) = 1 - 0,5987369 = 0,4012631$

(c) $P(X = 10) = \binom{10}{10}(0,05)^{10}(0,95)^{(10-10)} = (0,05)^{10}$

(d) $\mathrm{E}(X) = n\pi = 10 \times 0,05 = 0,5$

$\mathrm{Var}(X) = n\pi(1-\pi) = 10 \times 0,05 \times 0,95 = 0,475$

Exercício: Em uma amostra de cinco árvores, verificou-se a ocorrência de duas árvores com fuste de qualidade tipo 1 (fuste reto, cilíndrico, bem configurado e sem deterioração aparente). Sabe-se que a variável resposta "número de árvores com fuste de qualidade tipo 1", (X), segue a distribuição binomial. Obter a distribuição de probabilidade para a variável X.

$X_i = x_i$	$P(X_i = x_i)$
0	
1	
2	
3	
4	
5	
Total	1

Distribuição binomial

Exercício: Seja X a variável aleatória número de plantas com mutação em um total de n plantas irradiadas e $\pi = 0,0001$ a probabilidade de uma planta irradiada apresentar mutação. Calcular:

(a) A probabilidade de não aparecerem plantas com mutação em um total de 1000 plantas irradiadas;

(b) A probabilidade de aparecerem pelo menos uma planta com mutação em 1000 plantas irradiadas;

(c) A probabilidade de não aparecer planta com mutação em 2000 plantas irradiadas;

(d) A probabilidade de aparecerem pelo menos duas plantas com mutação em 2000 plantas irradiadas;

(e) O número médio esperado de plantas com mutação em 2000 plantas irradiadas;

(f) A variância esperada do número de plantas com mutação em 2000 plantas irradiadas;

(g) O número mínimo de plantas que devemos irradiar de modo que a probabilidade de aparecer ao menos uma planta com mutação seja maior do que ou igual a 0,90.

Solução

(a) $\pi = 0,0001 \quad n = 1000$
$P(X=0) = \binom{n}{x}\pi^x(1-\pi)^{n-x} =$
$\binom{1000}{0}(0,0001)^0(0,9999)^{(1000-0)} = 0,904833$

(b) $P(X \geq 1) = P(X=1) + P(X=2) + \ldots + P(X=1000) =$
$1 - P(X=0) = 1 - 0,904833 = 0,095167$

(c) $\pi = 0,0001 \quad n = 2000$
$P(X=0) = \binom{2000}{0}(0,0001)^0(0,9999)^{(2000-0)} = 0,818723$

(d) $P(X \geq 2) = P(X=2) + P(X=3) + \ldots + P(X=2000) = 1 -$
$P(X=0) - P(X=1) = 1 - \binom{2000}{0}(0,0001)^0(0,9999)^{(2000-0)} -$
$\binom{2000}{1}(0,0001)^1(0,9999)^{(2000-1)} = 1 - 0,818723 - 0,163761 =$
$0,017516$

(e) $\mathrm{E}(X) = n\pi = 2000 \times 0,0001 = 0,2$

(f) $\mathrm{Var}(X) = n\pi(1-\pi) = 2000 \times 0,0001 \times 0,9999 = 0,19998$

(g) $P(X \geq 1) = 1 - P(X=0) = 1 - \binom{n}{0}0,0001^0 \times 0,9999^n \geq 0,90$
$1 - 0,9 \geq 0,9999^n \Rightarrow \log(0,1) \geq n\log(0,9999) \Rightarrow n \geq 23024,7$

Distribuição de Poisson

Largamente utilizada quando se deseja analisar número de ocorrências de um evento de interesse, por unidade de tempo, comprimento, área ou volume. Exemplos:

- número de indivíduos por quadrante de 1 m^2;
- número de colônias de bactérias por 0,01 mm^2 de uma dada cultura, em uma plaqueta de laboratório;
- número de defeitos em 1000 m de tecido;
- número de acidentes em uma esquina movimentada e bem sinalizada, por dia;
- número de partículas radioativas emitidas numa unidade de tempo
- número de micronúcleos/1000 células.
- número de nematóides encontrados em amostras de solo.
- número de mortes por coronavirus, diariamente

Importante: Muito utilizada em estudos de dinâmica de populações e de entomologia.

Distribuição de Poisson

Definição

Se X representa o número de ocorrências de um evento no tempo ou no espaço e se X segue distribuição de Poisson de parâmetro λ, então, a função de probabilidade de X é dada por

$$P(x) = P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0, 1, 2, 3, \ldots,$$

em que λ é igual ao número de ocorrências do evento de interesse por unidade de tempo, distância, área, ..., etc e $e \approx 2,7183$.

Tem-se aqui, também, uma variável aleatória discreta com um número enumerável de valores.

Pressuposições

- independência dos eventos
- mesma taxa de ocorrência

Notação: $X \sim \mathrm{P}(\lambda)$.

Distribuição de Poisson

Média

A esperança de uma variável aleatória X com distribuição de Poisson com parâmetro λ é dada por:

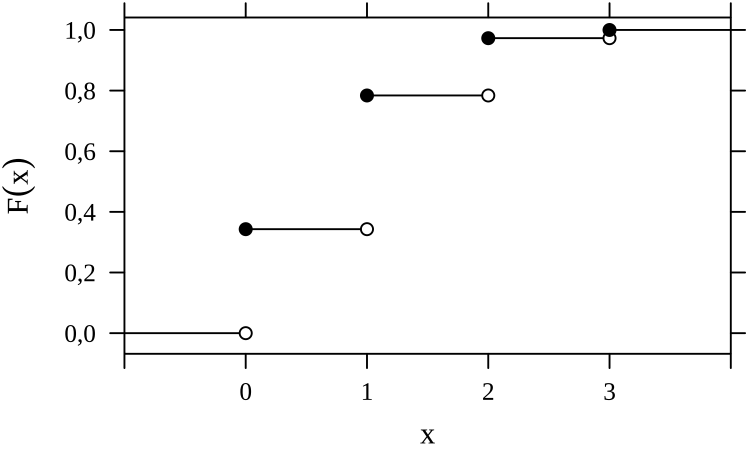
$$\mu_X = \mathrm{E}(X) = \lambda.$$

Variância

A variância de uma variável aleatória X com distribuição de Poisson com parâmetro λ é dada por:

$$\sigma_X^2 = \mathrm{Var}(X) = \lambda.$$

Representação gráfica



Distribuição de Poisson

Exemplo: Em um inventário florestal, verificou-se que há, em média, 20 árvores em parcelas de 0,25 ha. Sabe-se que a probabilidade de ocorrer uma árvore da família meliaceae é de 10% e que a variável resposta "número de árvores" dessa família segue a distribuição de Poisson. Calcular as probabilidades de ocorrência de:

(a) nenhuma árvore da família meliaceae

Dado que $\lambda = n\pi = 20 \times 0,10 = 2$

$P(X = 0) = P(0) = \frac{e^{-2}2^0}{0!} = 0,1353$

(b) apenas uma árvore da família meliaceae

$P(X = 1) = P(1) = \frac{e^{-2}2^1}{1!} = 0,1353 \times 2 = 0,2706$

(c) pelo menos três árvores da família meliaceae

$P(X \geq 3) = 1 - P(0) - P(1) - P(2) =$
$1 - 0,1353 - 0,2706 - \frac{e^{-2}2^2}{2!} = 0,3235$

(d) cinco árvores da família meliaceae

$P(X = 5) = P(5) = \frac{e^{-2}2^5}{5!} = \frac{0,1353 \times 32}{5 \times 4 \times 3 \times 2} = 0,0361$

Exemplo: A emissão de partículas radioativas tem sido modelada por meio de uma distribuição de Poisson, com o valor do parâmetro dependendo da fonte utilizada. Suponha que o número de partículas alfa, emitidas por minuto, seja uma variável aleatória seguindo o modelo de Poisson com taxa média de ocorrência de 5 emissões a cada minuto. Calcule a probabilidade de haver mais de duas emissões em um minuto.

Solução:

$\lambda = 5$

$P(X \geq 2) = 1 - P(0) - P(1) = 1 - \frac{e^{-5}5^0}{0!} - \frac{e^{-5}5^1}{1!} =$
$1 - 0,0067 - 0,0337 = 0,9596$

Distribuição de Poisson

Exercício: Os dados que se seguem referem-se a um estudo da distribuição da espécie *Primula simenses* selvagem em uma certa região florestal que foi dividida em 109 áreas. Acredita-se que a distribuição Poisson ajusta-se bem a esses dados.

$X_i = x_i$	$f(x_i)$	$f'(x_i)$	$P(X_i = x_i)$	$\hat{f}(x_i)$
0	26			
1	21			
2	23			
3	14			
4	11			
5	4			
6	5			
7	4			
8	1			
Total	$n = 109$			

x_i número de plantas por área
$f(x_i)$ número de áreas com x_i plantas (frequência observada)
$f'(x_i) = \frac{f(x_i)}{n}$ (frequência relativa)
$\hat{f}(x_i) = n \times P(X_i = x_i)$ (frequência esperada)

Distribuição de Poisson

(a) Qual a estimativa da probabilidade (frequência relativa observada) de se encontrarem pelo menos duas Primulas, escolhida uma área aleatoriamente?

(b) Calcule a média desses dados, usando

$$\bar{x} = \frac{1}{n} \sum_{i=0}^{k} x_i f(x_i), \quad n = \sum_{i=0}^{k} f(x_i)$$

(c) Faça $\lambda = \bar{x}$ e calcule as probabilidades $P(X_i = x_i)$ de acordo com

$$P(x_i) = P(X_i = x_i) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

(d) Calcule as frequências esperadas $\hat{f}(x_i) = n \times P(X_i = x_i)$

(e) Compare as frequências observadas com as frequências esperadas e responda a pergunta: "O modelo Poisson é realmente um bom modelo para ajustar-se a esses dados observados"? (Veremos um teste formal para isso, quando for estudada a distibuição χ^2)

Aproximação binomial pela Poisson

A distribuição de Poisson, $P(\lambda)$, com $\lambda = n\pi$ é uma boa aproximação para a distribuição binomial $B(n, \pi)$, quando π for pequeno e n for bastante grande, tal que $n\pi \leq 10$.

De fato, a distribuição Poisson é uma distribuição limite da binomial. Quando $n \to \infty$ e $\pi \to 0$ a distribuição binomial resulta na distribuição de Poisson com $\lambda = n\pi$.

Aproximação binomial pela Poisson

Exemplo: Seja X a variável número de plantas com mutação em um total de n plantas irradiadas e $\pi = 0,0001$ a probabilidade de uma planta irradiada apresentar mutação. Calcular:

(a) A probabilidade de não aparecer planta com mutação em 1000 plantas irradiadas;

(b) A probabilidade de aparecer pelo menos uma planta com mutação em 1000 plantas irradiadas;

(c) A probabilidade de não aparecer planta com mutação em 2000 irradiadas;

(d) A probabilidade de aparecerem ao menos duas plantas com mutação em 2000 plantas irradiadas;

(e) O número médio esperado de plantas com mutação em 2000 plantas irradiadas;

(f) A variância esperada do número de plantas com mutação em 2000 plantas irradiadas;

(g) O número mínimo de plantas que devem ser irradiadas de modo que a probabilidade de aparecer ao menos uma planta com mutação seja maior ou igual a 0,90.

Solução

(a) $\pi = 0,0001 \quad n = 1000 \quad \lambda = n \times \pi = 1000 \times 0,0001 = 0,1$
$P(X = 0) = \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-0,1} \times 0,1^0}{0!} = 0,904837$

(b) $P(X \geq 1) = P(X = 1) + P(X = 2) + \ldots + P(X = 1000) =$
$1 - P(X = 0) = 1 - 0,904837 = 0,095162$

(c) $\pi = 0,0001 \quad n = 2000 \quad \lambda = n \times \pi = 2000 \times 0,0001 = 0,2$
$P(X = 0) = \frac{e^{-0,2} \times 0,2^0}{0!} = 0,818731$

(d) $P(X \geq 2) = P(X = 2) + P(X = 3) + \ldots + P(X = 2000) =$
$1 - P(X = 0) - P(X = 1) = 1 - \frac{e^{-0,2} \times 0,2^0}{0!} - \frac{e^{-0,2} \times 0,2^1}{1!} =$
$1 - 0,818731 - 0,163746 = 0,017523$

(e) $\mathrm{E}(X) = \lambda = n\pi = 2000 \times 0,0001 = 0,2$

(f) $\mathrm{Var}(X) = \lambda = 0,2$

(f) $P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{e^{-n\pi}\lambda^0}{0!} = 1 - e^{-0,0001n} \geq 0,90$

$1 - 0,9 \geq e^{-0,0001n} \Rightarrow \log(0,1) \geq -0,0001n \Rightarrow n \geq 23025,85$