IND 1115 – Inferência Estatística – Semestre 2006.01
Teste 3 – 29/06/2005 – PROVA FINAL
GABARITO

PROBLEMA 1 (20 pontos)

O preço por litro da gasolina comum no Rio de Janeiro é uma variável aleatória Normal com média R$ 2.70 e desvio padrão R$ 0.15.

a) Qual a probabilidade da gasolina comum custar menos de R$ 2.60?

b) Quanto um posto deve cobrar pela gasolina para estar entre os 10% mais caros?

c) Quanto um posto deve cobrar pelo litro da gasolina para estar entre os 5% mais baratos?

Toma-se uma amostra de 9 postos de gasolina.

d) Qual a probabilidade do preço médio da gasolina na amostra ser menor que R$ 2.75 ?

Solução

X = preço da gasolina ~ $N(2.70, 0.15^2)$

a) $\Pr(X < 2.60) = \Pr\left(\frac{X - 2.70}{0.15} < \frac{2 - 2.70}{0.15}\right) = \Pr\left(Z < -\frac{0.10}{0.15}\right) = \Phi(-0.6667) = 1 - \Phi(0.6667) = 0.2525$

b) Para ser um dos 10% mais caros, a variável normalizada é z = 1.2816

Logo:

$$\frac{X - 2.70}{0.15} = 1.2816 \Leftrightarrow X = 2.70 + 0.15(1.2816) = R\$\,2.8922$$

c) Para estar entre os 5% mais baratos, a variável normalizada é z = -1.645 e então:

$$\frac{X - 2.70}{0.15} = -1.645 \Leftrightarrow X = 2.70 + 0.15(-1.645) = R\$\,2.4533$$

Agora considere uma amostra de 9 postos de gasolina. O preço médio é uma variável Normal com média R$ 2.70 minutos e variância $(0.15)^2/9$.

d) $\Pr(\overline{X} < 2.75) = \Pr\left(\frac{\overline{X} - 2.70}{0.15/3} < \frac{2.75 - 2.70}{0.15/3}\right) = \Pr\left(Z < \frac{0.05}{0.05}\right) = \Pr(Z < 1) = 0.8413$

PROBLEMA 2 (20 pontos)

Seja X uma variável aleatória contínua com densidade $f(x) = cx^2$ onde $0 < x < 2$.

a) Ache a constante c que faz de f(x) uma densidade.

b) Encontre a função de distribuição de X.

c) Ache a média, a variância e o desvio padrão de X.

d) Encontre um ponto m no intervalo (0,2) tal que $\Pr(X > m) = \Pr(X \le m) = 50\%$. Este ponto é a *mediana* da distribuição.

Solução

a) $\int_0^2 cx^2 dx = 1 \Rightarrow c\frac{x^3}{3}\Big|_0^2 = c\left(\frac{8}{3}\right) = 1 \Rightarrow c = \frac{3}{8}$

b) F(x) = 0 se x < 0 e F(x) = 1 se x > 2
Para X no intervalo [0,2]:

$$F(x) = \int_0^x 3\frac{u^2}{8} du = \frac{x^3}{8}$$

c) A média é:

$$E(X) = \int_0^{34} x\frac{3}{8}x^2 dx = \frac{3}{8}\left(\frac{x^4}{4}\right)\Big|_0^2 = \frac{3(16)}{8(4)} = \frac{6}{4} = 1.5$$

O segundo momento é:

$$E(X^2) = \int_0^4 x^2\frac{3}{8}x^2 dx = \frac{3}{8}\left(\frac{x^5}{5}\right)\Big|_0^2 = \frac{3(32)}{8(5)} = \frac{12}{5}$$

A variância é:

$$VAR(X) = \frac{12}{5} - (3/2)^2 = \frac{12}{5} - \frac{9}{4} = \frac{48-45}{20} = \frac{3}{20}$$

O desvio padrão é:
dp(X) = $\sqrt{(3/20)}$

d) A mediana é tal que: Pr(X ≤ m) = F(m) = 0.5. Logo, pelos resultados do item b):

$$\frac{m^3}{8} = \frac{1}{2} \Leftrightarrow m^3 = 4 \Leftrightarrow m = \sqrt[3]{4} = 1.5874$$

PROBLEMA 3 (20 pontos)

Sejam X_1 , X_2 , ... , X_{48} medidas independentes de uma certa experiência, e sejam Y_1 , Y_2 , ... , Y_{48} os X´s aproximados até o próximo inteiro. Seja $e_i = Y_i - X_i$, i =1, 2, ...48. Os e_i ‘s são iid Unif(-0.5, +0.5) .

Aproxime a seguinte probabilidade: $\Pr\left\{\left|\sum_{i=1}^{48} e_i\right| < 2.4\right\}$.

Solução

Pelas propriedades da distribuição Uniforme, cada e_i tem média 0 e variância 1/12. Logo:

$$E\left(\sum_{i=1}^{48} e_i\right) = 0$$

$$VAR\left(\sum_{i=1}^{48} e_i\right) = 48\left(\frac{1}{12}\right) = 4$$

Pelo teorema central do limite, o somatório dos e_i´s pode ser normalizado de forma a se tornar aproximadamente N(0,1).

$$\Pr\left\{\left|\sum_{i=1}^{48} e_i\right| < 2.4\right\} = \Pr\left(-2.4 < \sum_{i=1}^{48} e_i < +2.4\right) = \Pr\left(\frac{-2.4-0}{\sqrt{4}} < \frac{\sum_{i=1}^{48} e_i - 0}{\sqrt{4}} < \frac{+2.4-0}{\sqrt{4}}\right) \approx$$

$$\approx \Phi(1.2) - \Phi(-1.2) = 2\Phi(1.2) - 1 = 2(0.8849) - 1 = 0.7698$$

Problema 4 (20 pontos)

Sejam $X_1, X_2, ..., X_n$ iid Poisson(λ).

a) Encontre o estimador de máxima verossimilhança de λ.
b) O MLE de λ é não tendencioso?
c) O MLE de λ é consistente?
d) Encontre o estimador de máxima verossimilhança de Pr{ X_i = 0) = $e^{-\lambda}$. Verifique que este estimador de máxima verossimilhança é tendencioso . Encontre um estimador não tendencioso de $e^{-\lambda}$.
e) Ache o limite inferior de Cramer e Rao para λ.

Dica: Considere a função geradora de momentos dos X's.

Solução

a) A verossimilhança é:

$$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}(\lambda)^{x_i}}{x_i!} = \frac{e^{-n\lambda}(\lambda)^{\sum x_i}}{\prod_{i=1}^{n} x_i!} = \frac{e^{-n\lambda}(\lambda)^{n\overline{X}}}{\prod_{i=1}^{n} x_i!}$$

A log-verossimilhança é:
l(λ) = log L(λ) onde log indica o logaritmo natural (na base e)

$$l(\lambda) = -n.\lambda + n\overline{X}\log(\lambda) - \log\left(\prod_{i=1}^{n} x_i!\right)$$

Derivando-se a log-verossimilhança com relação a λ e igualando a derivada a zero produz o estimador de máxima verossimihança:

$$\frac{dl}{d\lambda} = -n + \frac{n\overline{X}}{\lambda} = 0 \Leftrightarrow \hat{\lambda} = \overline{X}$$

b) Note que $E(X_i) = \lambda$ para todo i = 1, 2, ...,n.

Mas:

$E(\hat{\lambda}) = E(\overline{X}) = \lambda$ e portanto o estimador de máxima verossimilhança é **não tendencioso** para λ .

c) Sabe-se que $VAR\left(\hat{\lambda}\right) = VAR\left(\overline{X}\right) = \frac{VAR\left(X_i\right)}{n} = \frac{\lambda}{n}$

Como o estimador de máxima verossimilhança é não tendencioso, seu erro quadrático médio é igual à sua variância, e portanto:

$$MSE\left(\hat{\lambda}\right) = VAR\left(\hat{\lambda}\right) = \frac{\lambda}{n}$$

O limite do MSE, quando n tende a infinito, é zero, e então o estimador de máxima verossimilhança é **consistente**.

d) Pelo princípio de invariância do estimador de máxima verossimilhança, o MLE de Pr{ X_i = 0) = $e^{-\lambda}$ é apenas: $\exp(-\overline{X})$.

Para avaliar se este estimador é não tendencioso, precisamos calcular:

$$E\left(\exp(-\overline{X})\right) = E\left(\exp\left(-\frac{\sum X_i}{n}\right)\right)$$

Mas, ΣX_i é uma variável Poisson com parâmetro n.λ, pois os X_i são independentes e identicamente distribuídos Poisson(λ).

Logo, a função geradora de momentos de U = ΣX_i é: $E(e^{tU}) = \exp\{ n.\lambda(e^t - 1) \}$.

Para verificar se o MLE de Pr{ X_i = 0) = $e^{-\lambda}$ é tendencioso ou não, precisamos calcular:

$$E\left(\exp(-\overline{X})\right) = E\left(\exp\left(-\frac{\sum X_i}{n}\right)\right) = E\left(\exp\left(-\frac{U}{n}\right)\right) = M_U\left(\frac{-1}{n}\right)$$

Onde M_U indica a função geradora de momentos de U.

$E\left(\exp(-\overline{X})\right) = M_U\left(\frac{-1}{n}\right) = \exp\left\{n.\lambda\left(e^{-1/n} - 1\right)\right\}$ que é diferente de exp(-λ).

Logo, o estimador de máxima verossimilhança de Pr{ X_i = 0) = $e^{-\lambda}$ é um estimador tendencioso.

e) Ache o Limite Inferior de Cramer e Rao para λ.

O Limite Inferior de Cramer e Rao (CRLB) é, neste caso:

CRLB = 1/I(λ)

Onde I(λ) é a informação de Fisher sobre o parâmetro λ.

$$I(\lambda) = -E\left(\frac{d^2 l}{d\lambda^2}\right)$$

Mas:

$$\frac{d^2 l}{d\lambda^2} = -\frac{n\overline{X}}{\lambda^2} = -\frac{U}{\lambda^2}$$

Então:

$$I(\lambda) = -E\left(\frac{d^2 l}{d\lambda^2}\right) = -E\left(\frac{-U}{\lambda^2}\right) = +\frac{1}{\lambda^2}E(U) = +\frac{1}{\lambda^2}(n\lambda) = \frac{n}{\lambda}$$

Logo:

$$CRLB = \frac{1}{I(\lambda)} = \frac{\lambda}{n}$$

Note que o estimador de máxima verossimilhança atinge o CRLB (ou seja, é um estimador eficiente), pois:

$$VAR(\hat{\lambda}) = \frac{\lambda}{n} = CRLB$$

Problema 5 (20 pontos)

Um computador gera 8 números aleatórios uniformemente distribuídos no intervalo (0,1).

a) Calcule a probabilidade de que o menor destes números será menor que 0.1.

b) Calcule o valor esperado do menor destes números.

c) Encontre a densidade do maior destes 8 números.

d) Encontre o valor esperado do maior destes 8 números.

e) Calcule a probabilidade de que o maior destes números exceda 0.8.

Dica: você pode citar resultados dos slides, ao invés de demonstrar explicitamente todos os passos necessários aqui.

Solução

Teorema
Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias independentes com densidade Unif(0,1). Seja Y_r o r-ésimo maior número dentre os valores observados de $X_1, X_2, \ldots, X_n$. Então Y_r tem densidade Beta com parâmetros r e n –r + 1.

Solução

a) A densidade do menor dos 8 números é uma Beta com parâmetros 1 e 8. Isto é, se Y denota este número temos:

$$f(y) = \frac{\Gamma(9)}{\Gamma(1).\Gamma(8)} y^{1-1}(1-y)^{8-1} = \frac{8!}{0!7!}(1-y)^7 = 8(1-y)^7 \quad \text{onde } 0<y<1$$

A probabilidade deste número ser menor que 0.1 é:

Pr{ Y < 0.1} = $\int_0^{0.1} 8(1-y)^7\, dy$

Faça a mudança de variável: t = 1 - y $\Rightarrow$ dt = -dy e se y$\to$ 0.1, t $\to$ 0.9 e se y$\to$ 0, t $\to$ 1. Logo:

Pr{ Y < 0.1} = $\int_1^{0.9} 8t^7(-dt) = 8\int_{0.9}^{1} t^7 dt = t^8 \Big|_{0.9}^{1} = 1-(0.9)^8 = 0.5695$

b) Calcule o valor esperado do menor destes números.

Y é Beta (1,8) e portanto seu valor esperado é 1/(1+8) = 1/9 = 0.1111

c) A densidade do maior destes números é, pelo teorema, Beta(8,1).

d) Seja W o maior destes 8 números. Então E(W) = 8/(1+8) = 8/9 = 0.8888

e) Pr(W > 0.8) = ?
A densidade de W é:

$$f(w) = \frac{\Gamma(9)}{\Gamma(1).\Gamma(8)} w^{8-1}(1-w)^{1-1} = \frac{8!}{0!7!} w^7 = 8w^7 \quad \text{onde } 0<w<1$$

Pr(W > 0.8) = $\int_{0.8}^{1} 8w^7 dw = w^8 \Big|_{0.8}^{1} = 1-(0.8)^8 = 0.8322$

Tabela – Função de Distribuição N(0,1)

z	Φ(z)		z	Φ (z)		z	Φ (z)
0.0000	50.00%		0.9800	83.65%		2.0125	97.79%
0.0200	50.80%		0.9900	83.89%		2.0200	97.83%
0.0300	51.20%		1.0000	84.13%		2.0300	97.88%
0.0400	51.60%		1.0100	84.38%		2.0400	97.93%
0.0500	51.99%		1.0167	84.54%		2.0412	97.94%
0.1000	53.98%		1.0250	84.73%		2.0500	97.98%
0.1500	55.96%		1.0500	85.31%		2.1000	98.21%
0.1667	56.62%		1.0553	85.44%		2.1500	98.42%
0.2000	57.93%		1.0800	85.99%		2.2000	98.61%
0.2236	58.85%		1.1000	86.43%		2.2361	98.73%
0.2500	59.87%		1.1180	86.82%		2.3000	98.93%
0.3000	61.79%		1.1475	87.44%		2.3263	99.00%
0.3015	61.85%		1.1500	87.49%		2.3333	99.02%
0.3333	63.06%		1.1553	87.60%		2.4000	99.18%
0.3475	63.59%		1.2000	88.49%		2.5000	99.38%
0.3492	63.65%		1.2060	88.61%		2.5500	99.46%
0.3500	63.68%		1.2200	88.88%		2.5628	99.48%
0.4000	65.54%		1.2500	89.44%		2.6000	99.53%
0.4167	66.16%		1.2700	89.79%		2.6500	99.60%
0.4307	66.67%		1.2816	90.00%		2.6667	99.62%
0.4500	67.36%		1.3000	90.32%		2.6833	99.64%
0.5000	69.15%		1.3333	90.88%		2.7000	99.65%
0.5500	70.88%		1.3750	91.54%		2.7500	99.70%
0.5774	71.81%		1.4000	91.92%		2.8000	99.74%
0.6000	72.57%		1.4468	92.60%		2.9000	99.81%
0.6250	73.40%		1.4500	92.65%		2.9500	99.84%
0.6500	74.22%		1.5000	93.32%		3.0000	99.87%
0.6667	74.75%		1.5500	93.94%		3.1000	99.90%
0.6708	74.88%		1.5811	94.31%		3.1500	99.92%
0.7000	75.80%		1.6000	94.52%		3.2000	99.93%
0.7500	77.34%		1.6450	95.00%			
0.8000	78.81%		1.6667	95.22%			
0.8333	79.77%		1.7000	95.54%			
0.8500	80.23%		1.8000	96.41%			
0.8666	80.69%		1.8500	96.78%			
0.8944	81.45%		1.9000	97.13%			
0.9000	81.59%		1.9500	97.44%			
0.9167	82.03%		1.9600	97.50%			
0.9500	82.89%		1.9800	97.61%			
0.9600	83.15%		1.9900	97.67%			
0.9700	83.40%		2.0000	97.72%			
0.9750	83.52%		2.0100	97.78%			