IND 1115 – Inferência Estatística – Semestre 2006.01
Teste 2 – 20/06/2005
GABARITO

PROBLEMA 1 (16 pontos)

Considere uma variável Binomial com n = 36 e p = ½.

a) Calcule exatamente:

x	**Pr(X = x)**
17	0.1251
18	0.1321
19	0.1251
Pr(17 ≤ X ≤19)	**0.3823**

b) Calcule usando o Teorema de DeMoivre e Laplace **com** correção de continuidade.

x	**Pr(X = x)**
17	0.1253
18	0.1324
19	0.1253
Pr(17 ≤ X ≤19)	**0.3830**

Seja X a variável Binomial. Então, pelo teorema de DeMoivre e Laplace:

$$\frac{X-18}{\sqrt{36(1/2)(1/2)}}=\frac{X-18}{3}$$ é aproximadamente N(0,1).

Os valores individuais só podem ser calculados através da correção de continuidade como abaixo:

$$\Pr(X=17)=\Pr(16.5<X<17.5)=\Pr\left(\frac{16.5-18}{3}<\frac{X-18}{3}<\frac{17.5-18}{3}\right)=$$

$$=\Phi\left(\frac{-0.5}{3}\right)-\Phi\left(\frac{-1.5}{3}\right)=\Phi(-0.1667)-\Phi(-0.5)=$$

$$=\left[1-\Phi(+0.1667)\right]-\left[1-\Phi(+0.5)\right]=$$

$$=\Phi(+0.5)-\Phi(+0.1667)=0.6914-0.5662=0.1253$$

$$\Pr(X = 18) = \Pr\left(17.5 < X < 18.5\right) = \Pr\left(\frac{17.5-18}{3} < \frac{X-18}{3} < \frac{18.5-18}{3}\right) =$$

$$= \Phi\left(\frac{0.5}{3}\right) - \Phi\left(\frac{-0.5}{3}\right) = \Phi\left(0.1667\right) - \Phi\left(-0.1667\right) =$$

$$= \Phi\left(0.1667\right) - \left[1 - \Phi\left(+0.1667\right)\right] = 2.\Phi\left(0.1667\right) - 1 =$$

$$= 2(0.5662) - 1 = 0.1324$$

$$\Pr(X = 19) = \Pr\left(18.5 < X < 19.5\right) = \Pr\left(\frac{18.5-18}{3} < \frac{X-18}{3} < \frac{19.5-18}{3}\right) =$$

$$= \Phi\left(\frac{1.5}{3}\right) - \Phi\left(\frac{0.5}{3}\right) = \Phi\left(0.5\right) - \Phi\left(0.1667\right) =$$

$$= 0.6915 - 0.5662 = 0.1253$$

A soma destes três valores fornece o valor aproximado de Pr(17 ≤ X ≤19), que é: 0.3830

PROBLEMA 2 (16 pontos)

Seja X uma variável aleatória Qui-quadrado com 72 graus de liberdade. Use o teorema central do limite para aproximar:

a) Pr(X < 90)
b) Pr(X > 60)

Solução

X pode ser encarado como uma soma de 72 variáveis independentes, cada uma com densidade Qui-quadrado com 1 grau de liberdade, e portanto podemos usar o teorema central do limite.

A média de X é 72 e sua variância é 144. Logo:

$\frac{X-72}{\sqrt{144}} = \frac{X-72}{12}$ é aproximadamente N(0,1).

a) $\Pr\left(X < 90\right) = \Pr\left(\frac{X-72}{12} < \frac{90-72}{12}\right) = \Phi\left(1.5\right) = 0.9332$

b) $\Pr\left(X > 60\right) = \Pr\left(\frac{X-72}{12} > \frac{60-72}{12}\right) = 1 - \Phi\left(-1\right) = \Phi(+1) = 0.8413$

PROBLEMA 3 (20 pontos)

Considere uma amostra aleatória de tamanho n da distribuição Uniforme(-θ, + θ).

a) Suponha que você observa as seguintes amostras independentes (que já foram ordenadas para facilitar o seu trabalho):

Amostra 1	Amostra 2	Amostra 3
-1,60	-1,94	-1,12
-0,47	-1,82	-0,86
0,39	-1,45	-0,63
1,54	-0,37	0,45
1,60	1,02	1,34
1,83	1,45	1,78

Intuitivamente, qual o valor do estimador de máxima verossimilhança de θ na amostra 1? E na amostra 2? E na amostra 3?

b) Encontre o estimador de máxima verossimilhança de θ.

Solução

a) Intuitivamente as observações mais extremas são aquelas que estão trazendo maior informação sobre o parâmetro desconhecido θ. No caso da amostra 1, o MLE é, portanto, 1.83. Na amostra 2, o mínimo é mais extremo que o máximo, e o MLE é 1.94. Na amostra 3 o MLE é 1.78.

b) A densidade de cada X_i é:

$$f(x,\theta) = \frac{1}{\theta - (-\theta)} = \frac{1}{2\theta} \quad \text{se } -\theta \le x \le \theta$$

A verossimilhança é:

$$L(\theta) = \frac{1}{(2\theta)^n} \quad \text{se } -\theta \le X_{(1)} \le \ldots \le X_{(n)} \le \theta$$

Logo, a verossimilhança é máxima se θ é pequeno. Mas, para que a verossimilhança seja não nula, θ tem que, simultaneamente, satisfazer as duas equações a seguir:

$$\theta \ge X_{(n)} = \max(X_1, X_2, .., X_n) \text{ e } -\theta \le X_{(1)} = \min(X_1, X_2, .., X_n)$$

Ou seja:

$$\theta \ge X_{(n)} \text{ e } \theta \ge -X_{(1)}$$

Isto e´´:

$$\theta \ge \max\left(X_{(n)}, -X_{(1)}\right)$$

O menor valor que satisfaz estas duas equações é:

$\hat{\theta} = \max\left(X_{(n)}, -X_{(1)}\right)$, que é o MLE de θ.

Problema 4 (30 pontos)

Considere uma amostra aleatória de tamanho n da distribuição Bernoulli(p).

a) Encontre o estimador de máxima verossimilhança de p.
b) O MLE é não tendencioso para p? É consistente para p?
c) Ache a informação de Fisher.
d) Encontre o limite inferior de Cramér e Rao.
e) O estimador de máxima verossimilhança encontrado é um estimador eficiente de p?

JUSTIFIQUE AS SUAS RESPOSTAS, ESPECIALMENTE NOS ITENS b) E e).

Solução

a) A função de probabilidade de cada X_i é:

$$f(x, p = p^x (1-p)^{1-x} \quad \text{onde x = 0,1}$$

A verossimilhança é:

$$L(p) = p^{\sum_{i=1}^n X_i} (1-p)^{n-\sum_{i=1}^n X_i}$$

A log-verossimilhança é:

$$l(p) = \sum_{i=1}^n X_i \left\{\log(p)\right\} + \left[n - \sum_{i=1}^n X_i\right] \log(1-p)$$

Derivando a log-verossimilhança com relação a p e igualando a zero:

$$\frac{\partial l}{\partial p} = \frac{\sum_{i=1}^n X_i}{p} - \frac{n - \sum_{i=1}^n X_i}{1-p} = 0 \Leftrightarrow \frac{\sum_{i=1}^n X_i}{p} = \frac{n - \sum_{i=1}^n X_i}{1-p}$$

$$\Leftrightarrow np - p\sum_{i=1}^n X_i = \sum_{i=1}^n X_i - p\sum_{i=1}^n X_i \Leftrightarrow np = \sum_{i=1}^n X_i$$

$$\Leftrightarrow \hat{p} = \frac{1}{n}\sum_{i=1}^n X_i = \overline{X}$$

é o MLE de p.

b) Note que:

$$\sum_{i=1}^{n} X_i \sim \text{Bin(n,p)} \Rightarrow \text{E}\left(\sum_{i=1}^{n} X_i\right) = np \quad \text{e}\, VAR\left(\sum_{i=1}^{n} X_i\right) = npq$$

$$Tambem:$$

$$E\left(\bar{X}\right) = p \quad \text{e VAR}\left(\bar{X}\right) = \frac{pq}{n}$$

Logo, o MLE é não tendencioso para p. Também é consistente para p, pois o limite de sua variância é zero quando n tende a infinito.

c) A informação de Fisher é:

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

Mas:

$$\frac{\partial^2 l}{\partial p^2} = \frac{\partial}{\partial p}\left\{\frac{\sum X_i}{p} - \frac{n - \sum X_i}{1-p}\right\} = -\frac{\sum X_i}{p^2} - \frac{n - \sum X_i}{\left(1-p\right)^2}$$

$$\Rightarrow I(p) = -E\left(\frac{\partial^2 l}{\partial p^2}\right) = -E\left(-\frac{\sum X_i}{p^2} - \frac{n - \sum X_i}{\left(1-p\right)^2}\right) =$$

$$= \frac{np}{p^2} + \frac{n - np}{\left(1-p\right)^2} = \frac{n}{p} + \frac{n(q)}{q^2} = \frac{n}{p} + \frac{n}{q} = \frac{np + nq}{pq} = \frac{n}{pq} = \frac{1}{pq/n} = \frac{1}{VAR(\bar{X})}$$

d) O limite inferior de Cramer e Rao para p é:

$$CRLB = \frac{1}{I(p)} = \frac{pq}{n} = VAR\left(\bar{X}\right)$$

e) Pelo item anterior, vemos que a variância do MLE é igual ao limite inferior de Cramer e Rao. Além disso, ele é um estimador não tendencioso. Logo, é um estimador eficiente de p.

Problema 5 (18 pontos)

Seja X_1 , X_2 , ... , X_n uma amostra aleatória da densidade N(1,θ).

Encontre o estimador de máxima verossimilhança de θ.

Solução

A densidade de cada X_i é:

$$f(x,\theta) = \frac{1}{\sqrt{2\pi.\theta}} \exp\left\{ \frac{-1}{2\theta} (x-1)^2 \right\}$$

A verossimilhança é:

$$L(\theta) = (2.\pi.\theta)^{-n/2} \exp\left\{ \frac{-\sum_{i=1}^{n} (X_i - 1)^2}{2\theta} \right\}$$

A log-verossimilhança é:

$$l(\theta) = -\frac{n}{2}\log(2\pi.\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(X_i - 1)^2 \Leftrightarrow$$

$$\frac{\partial l}{\partial \theta} = \frac{-n}{2\theta} + \frac{\sum_{i=1}^{n}(X_i - 1)^2}{2\theta^2} = 0 \Leftrightarrow \frac{\sum_{i=1}^{n}(X_i - 1)^2}{2\theta^2} = \frac{n}{2\theta}$$

$$\Leftrightarrow \frac{\sum_{i=1}^{n}(X_i - 1)^2}{n} = \hat{\theta}$$

é o estimador de máxima verossimilhança de θ.

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%			