IND 1115 – Inferência Estatística – Semestre 2005.01 – turma A
Teste 2 – 09/06/2005
GABARITO

PROBLEMA 1 (20 pontos)

Em cada questão abaixo, indique se a afirmativa é **verdadeira** ou **falsa** (marque um X na alternativa correta). Não é necessário justificar a sua resposta.

		Verdadeiro	Falso
1	A correção de continuidade existe no teorema de DeMoivre e Laplace e nas outras aplicações do Teorema Central do Limite		XXX
2	Em Estatística os parâmetros de uma densidade são conhecidos.		XXX
3	A densidade t de Student é assimétrica em torno de zero.		XXX
4	Na estimação pontual buscamos encontrar um intervalo que contenha a variável aleatória θ com uma probabilidade especificada.		XXX
5	Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma densidade com média a e variância b. Então a média amostral tem também média a e variância b.		XXX
6	O princípio de invariância do MLE garante que, se T é não tendencioso para θ, então g(T) é não tendencioso para $g(\theta)$.		XXX
7	Um estimador é consistente se sua variância tende a zero quando o tamanho da amostra cresce indefinidamente.		XXX
8	Existem estimadores não tendenciosos com variância menor que o limite inferior de Cramér e Rao.		XXX
9	O método de momentos fornece estimadores únicos.		XXX
10	O método de momentos fornece estimadores iguais aos de máxima verossimilhança.		XXX

Problema 2 (10 pontos)

Aproxime, com base no Teorema Central do Limite, as seguintes probabilidades:

$$a)\Pr\left(\chi^2_{25} \le 27\right) = \Pr\left(\frac{\chi^2_{25} - 25}{\sqrt{50}} \le \frac{27-25}{\sqrt{50}}\right) \approx \Phi\left(\frac{27-25}{\sqrt{50}}\right) = \Phi(0.2828) = 0.6114$$

$$b)\Pr\left(\chi^2_{36} > 48\right) = \Pr\left(\frac{\chi^2_{36} - 36}{\sqrt{72}} \le \frac{48-36}{\sqrt{72}}\right) \approx \Phi\left(\frac{48-36}{\sqrt{72}}\right) = \Phi(1.4142) = 0.9214$$

Problema 3 (30 pontos)

Considere uma amostra aleatória de tamanho n da distribuição Geométrica(p), ou seja, a função de probabilidade é dada por:

$$\Pr(X = x) = f(x) = q^{x-1}p = (1-p)^{x-1}p \quad \text{para } x = 1,2,3.....$$

a) Encontre o estimador de máxima verossimilhança de p.

b) Encontre o MLE de Pr(X > 2).

c) Encontre a informação de Fisher.

Dica: Série Geométrica

$$\sum_{x=1}^{\infty} a^x = \frac{a}{1-a} \qquad \text{se } |a| < 1$$

Dica 2: Se X tem distribuição Geométrica(p) então E(X) = 1/p

Solução

a) $L(p) = \prod_{i=1}^{n} (1-p)^{x_i - 1} p = p^n (1-p)^{-n} (1-p)^{\sum x_i}$

$$L(p) = p^n (1-p)^{n\overline{X} - n}$$

$$l(p) = \log L(p) = n\log(p) + (n\overline{X} - n).\log(1-p)$$

$$\frac{dl}{dp} = 0 \Leftrightarrow \frac{n}{p} + \frac{(n\overline{X} - n)(-1)}{1-p} = 0 \Leftrightarrow \frac{n}{p} = \frac{(n\overline{X} - n)}{1-p} \Leftrightarrow \frac{1}{p} = \frac{(\overline{X} - 1)}{1-p} \Leftrightarrow 1 - p = p\overline{X} - p \Leftrightarrow 1 = p\overline{X} \Leftrightarrow \hat{p} = \frac{1}{\overline{X}}$$

$$\hat{p} = \frac{1}{\overline{X}} = \frac{n}{\sum X_i}$$

b) Pr(X > 2) = Pr(X = 3) + Pr(X = 4) + ... = 1 – Pr(X = 1) – Pr(X = 2) = 1 – p – pq. Pelo princípio de invariância do MLE, o estimador de máxima verossimilhança de Pr(X > 2) é:

$$1 - \hat{p} - \hat{p}(1-\hat{p}) = 1 - \frac{1}{\overline{X}} - \left(\frac{1}{\overline{X}}\right)\left(1 - \frac{1}{\overline{X}}\right)$$

c) A Informação de Fisher pode ser encontrada a partir de:

$$\frac{d^2 l}{dp^2} = \frac{d}{dp}\left\{\frac{n}{p} - \frac{\left(n\overline{X} - n\right)}{1-p}\right\} = \frac{-n}{p^2} - \frac{\left(n\overline{X} - n\right)}{\left(1-p\right)^2}$$

$$I(p) = -E\left\{\frac{d^2 l}{dp^2}\right\} = -E\left\{\frac{-n}{p^2} - \frac{\left(n\overline{X} - n\right)}{\left(1-p\right)^2}\right\} = \frac{n}{p^2} + \frac{E\left(n\overline{X} - n\right)}{\left(1-p\right)^2} = \frac{n}{p^2} + \frac{\left(n/p - n\right)}{\left(q\right)^2} = \frac{n}{p^2} + \frac{n\left(1-p\right)}{p\left(q\right)^2} = \frac{n}{p^2} + \frac{nq}{p\left(q\right)^2} =$$

$$= \frac{n}{p^2} + \frac{n}{pq} = \frac{n}{p}\left\{\frac{1}{p} + \frac{1}{q}\right\} = \frac{n}{p}\left\{\frac{1}{pq}\right\} = \frac{n}{p^2 q}$$

Problema 4 (40 pontos)

Sejam $X_1, X_2, ..., X_n$ iid N(μ, 4). Ou seja, a densidade de cada X_i é:

$$f(x_i, \mu) = \frac{1}{\sqrt{2(4)\pi}} \exp\left\{-\frac{1}{2(4)}\left[x_i - \mu\right]^2\right\}$$ para X_i um número real.

a) Encontre o MLE de μ.
b) Mostre que o MLE é não tendencioso.
c) Mostre que o MLE é consistente.
d) Calcule a informação de Fisher.
e) Calcule o limite inferior de Cramer e Rao. O MLE é um estimador eficiente?
f) Encontre um estimador por método de momentos de μ.
g) Encontre, a partir da média amostral, um estimador não tendencioso de μ^2.

Solução

a) $$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2(4)\pi}} \exp\left\{-\frac{1}{2(4)}\left[x_i - \mu\right]^2\right\} = \left(8\pi\right)^{-n/2} \exp\left\{-\frac{1}{8}\sum_{i=1}^{n}\left[x_i - \mu\right]^2\right\}$$

$$l(\mu) = \log L(\mu) = -\frac{n}{2}\log\left(8\pi\right) - \frac{1}{8}\sum_{i=1}^{n}\left[x_i - \mu\right]^2 = -\frac{n}{2}\log\left(8\pi\right) - \frac{1}{8}\sum_{i=1}^{n}\left[x_i^2 - 2\mu . x_i + \mu^2\right] =$$

$$= -\frac{n}{2}\log\left(8\pi\right) - \frac{1}{8}\left\{\sum_i x_i^2 - 2\mu\sum_i x_i + n\mu^2\right\}$$

$$\frac{dl}{d\mu} = 0 \Leftrightarrow -\frac{1}{8}\frac{d}{d\mu}\left\{\sum_i x_i^2 - 2\mu\sum_i x_i + n\mu^2\right\} = 0 \Leftrightarrow -2\sum_i x_i + 2n\mu = 0 \Leftrightarrow -n\overline{X} + n\mu = 0 \Leftrightarrow \hat{\mu} = \overline{X}$$

b) O MLE é não tendencioso pois $\overline{X}$ é Normal com média μ e variância 4/n.

c) O erro quadrático médio do MLE é apenas a sua variância neste caso, pois o estimador é não tendencioso. Do item acima, a variância do MLE é 4/n, que tende a zero quando n tende a infinito. Logo, o erro quadrático médio vai a zero quando n tende a infinito e o estimador é consistente.

d) A informação de Fisher é:

$$I(\mu) = -E\left\{\frac{d^2 l}{d\mu^2}\right\} = -E\left\{\frac{d}{d\mu}\left\{-\frac{1}{8}\frac{d}{d\mu}\left\{\sum_i x_i^2 - 2\mu\sum_i x_i + n\mu^2\right\}\right\}\right\} = +\frac{1}{8}\left(2n\right) = \frac{n}{4} = \frac{1}{VAR\left(\overline{X}\right)}$$

e) Calcule o limite inferior de Cramer e Rao. O MLE é um estimador eficiente?

O limite inferior de Cramer e Rao é:

$CRLB = \frac{1}{I(\mu)} = \frac{4}{n} = VAR(\overline{X})$ e portanto o MLE atinge o limite inferior de Cramer e Rao e é um estimador eficiente.

f) Um estimador por método de momentos de μ é obtido fazendo-se:

$\frac{1}{n}\sum_i X_i = \overline{X} = E(X_i) = \mu \Leftrightarrow \tilde{\mu} = \overline{X}$ e coincide com o estimador de máxima verossimilhança.

g) Um estimador tentativo de μ^2 é $\overline{X}^2$. Mas, $E(\overline{X}) = \mu$ e $\mathrm{VAR}(\overline{X}) = \frac{4}{n}$

Logo, $VAR(\overline{X}) = E(\overline{X}^2) - \mu^2 = \frac{4}{n} \Leftrightarrow E(\overline{X}^2) = \mu^2 + \frac{4}{n}$ e um estimador não tendencioso de μ^2 pode ser obtido fazendo-se: $\overline{X}^2 - \frac{4}{n}$.

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%
0,2000	57,93%		1,0553	85,44%		2,2000	98,61%
0,2236	58,85%		1,1000	86,43%		2,2361	98,73%
0,2828	61,14%		1,1180	86,82%		2,3000	98,93%
0,2500	59,87%		1,1475	87,44%		2,3263	99,00%
0,3000	61,79%		1,1500	87,49%		2,3333	99,02%
0,3015	61,85%		1,1553	87,60%		2,4000	99,18%
0,3333	63,06%		1,2000	88,49%		2,5000	99,38%
0,3475	63,59%		1,2060	88,61%		2,5500	99,46%
0,3492	63,65%		1,2200	88,88%		2,5628	99,48%
0,3500	63,68%		1,2500	89,44%		2,6000	99,53%
0,4000	65,54%		1.2700	89.79%		2,6500	99,60%
0.4167	66.16%		1,2816	90,00%		2,6667	99,62%
0,4307	66,67%		1,3000	90,32%		2,6833	99,64%
0,4500	67,36%		1,3333	90,88%		2,7000	99,65%
0,5000	69,15%		1,3750	91.54%		2.7500	99.70%
0,5500	70,88%		1,4000	91,92%		2,8000	99,74%
0,5774	71.81%		1,4142	92,14%		2,9000	99,81%
0,6000	72,57%		1,4500	92,65%		2,9500	99,84%
0,6250	73,40%		1,5000	93,32%		3,0000	99,87%
0,6500	74,22%		1,5500	93,94%		3,1000	99,90%
0,6667	74,75%		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%			