IND 1115 – Inferência Estatística – Semestre 2004.02
Teste 3 – 09/12/2004
GABARITO

Problema 1 (20 pontos)

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma densidade Uniforme(θ, 2) onde θ é desconhecido.
a) Encontre o MLE (estimador de máxima verossimilhança) de θ.
b) Encontre um estimador por método de momentos de θ.
c) O MLE de θ é não tendencioso?
d) Coleta-se uma amostra de tamanho 10 da Uniforme(θ, 2). Os valores observados são:
1,54 0,20 0,42 0,60 0,94
1,80 0,06 1,60 1,44 0,56
Compare o MLE e o estimador por método dos momentos de θ baseados nesta amostra.

Solução

A densidade de cada X_i é:

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

A verossimilhança é:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{2-\theta} = \frac{1}{(2-\theta)^n} \quad \text{se } \theta \le x_{(1)} = \min(x_1, x_2, \ldots, x_n) \le \ldots \le x_{(n)} = \max(x_1, x_2, \ldots, x_n)$$

Á medida que θ se aproxima de 2, a verossimilhança cresce, mas existe um valor máximo para θ, que é dado por $x_{(1)}$, o mínimo da amostra.Logo, o máximo da verossimilhança ocorre quando $\theta = x_{(1)}$.

b) Um estimador de θ por método de momentos iguala a média amostral à média da distribuição, que é $(\theta+2)/2$. Logo:

$$\overline{X} = \frac{\theta+2}{2} \Leftrightarrow \tilde{\theta} = 2\overline{X} - 2$$ é o estimador por método de momentos.

A função de distribuição do MLE é:

$$F(x) = \Pr\left(X_{(1)} \le x\right) = 1 - \Pr\left(X_{(1)} > x\right) = 1 - \Pr\left(X_1 > x, \ldots, X_n > x\right) =$$

$$= 1 - \prod_{i=1}^{n} \Pr\left(X_i > x\right) = 1 - \left\{\int_x^2 \frac{1}{2-\theta} du\right\}^n = 1 - \left(\frac{2-x}{2-\theta}\right)^n$$

A densidade do MLE é apenas a derivada desta função de distribuição, ou seja:

$$g(x) = \frac{dF}{dx} = -n\left(\frac{2-x}{2-\theta}\right)^{n-1}(-1) = n\left(\frac{2-x}{2-\theta}\right)^{n-1} \quad \text{para x } \in [\theta, 2]$$

A média do MLE é:

$$E\left(X_{(1)}\right)=\int_{\theta}^{2} nx\left(\frac{2-x}{2-\theta}\right)^{n-1} dx$$

Faça a transformação de variáveis: u = 2-x ⇒ x = 2- u , du = - dx

$$E\left(X_{(1)}\right)=\int_{0}^{2-\theta} n(2-u)\left(\frac{u}{2-\theta}\right)^{n-1} du=\frac{n}{(2-\theta)^{n-1}}\int_{0}^{2-\theta}(2-u)(u)^{n-1} du=$$

$$=\frac{n}{(2-\theta)^{n-1}}\left\{\frac{2u^{n}}{n}-\frac{u^{n+1}}{n+1}\right\}\Bigg|_{0}^{2-\theta}=\frac{n}{(2-\theta)^{n-1}}\left\{\frac{2(2-\theta)^{n}}{n}-\frac{(2-\theta)^{n+1}}{n+1}\right\}=$$

$$=\frac{(2-\theta)(2+\theta n)}{n+1}$$

Claramente o estimador de máxima verossimilhança é tendencioso para θ.

A média amostral é: 0.9160

O estimador por método de momentos baseado na amostra é: 2(0.916) – 2 = -0.168
O estimador de máxima verossimilhança é apenas $x_{(1)}$ = mínimo = 0.06

Problema 2 (20 pontos)

Sejam X_1 , X_2 , ... , X_n iid $N(0, \sigma^2)$.
a) Qual a média e variância de X_i^2?
b) Como aproximar $\Pr\{ X_1^2 + X_2^2 + ...+ X_n^2 \leq x)$ em termos de $\Phi(.)$?

Solução

$\frac{X_i^2}{\sigma^2}$ tem densidade qui-quadrado com 1 grau de liberdade para i =1,2,...n

Logo, sua média e variância são:

$$E\left(\frac{X_i^2}{\sigma^2}\right)=1 \Rightarrow \mathrm{E}\left(X_i^2\right)=\sigma^2 \text{ e } \mathrm{VAR}\left(\frac{X_i^2}{\sigma^2}\right)=2 \Rightarrow \quad \mathrm{VAR}\left(X_i^2\right)=2\sigma^4$$

Então, a soma dos X_i^2 ´s tem média e variância:

$$E\left(\sum_{i=1}^{n} X_i^2\right)=n\sigma^2$$

$$VAR\left(\sum_{i=1}^{n} X_i^2\right)=2n\sigma^4$$

Pelo Teorema Central do Limite, se n for grande podemos normalizar esta soma de tal forma que ela seja aproximadamente N(0,1).

Seja T = $X_1^2 + X_2^2 + ...+ X_n^2$ então:

$$Z = \frac{T - n\sigma^2}{\sqrt{2n\sigma^4}} = \frac{\sum_{i=1}^{n} X_i^2 - n\sigma^2}{\sqrt{2n\sigma^4}} \approx N(0,1)$$

$$\Pr(T \le x) = \Pr\left(\frac{T - n\sigma^2}{\sqrt{2n\sigma^4}} \le \frac{x - n\sigma^2}{\sqrt{2n\sigma^4}}\right) \approx \Phi\left(\frac{x - n\sigma^2}{\sqrt{2n\sigma^4}}\right)$$ é a aproximação desejada.

Problema 3 (20 pontos)

Sejam X_1 , X_2 , ... , X_{48} medidas independentes de uma certa experiência, e Y_1 , Y_2 , ... , Y_{48} 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)

Seja X uma variável aleatória com função de probabilidade Geométrica com probabilidade p, isto é: $\Pr(X = x) = (1-p)^{x-1} p$ onde $x = 1,2,....$

Encontre a função geradora de momentos de X e, a partir dela, a média de X.

Solução

A função geradora de momentos de X é:

$$E\left(e^{tX}\right)=\sum_{i=1}^{\infty}e^{tx}\left(1-p\right)^{x-1}p=p\left(1-p\right)^{-1}\sum_{i=1}^{\infty}\left\{e^{tx}\left(1-p\right)^{x}\right\}=p\left(q\right)^{-1}\sum_{i=1}^{\infty}\left\{\left(qe^{t}\right)^{x}\right\}=$$

$$=\frac{p}{q}\left\{\frac{qe^{t}}{1-qe^{t}}\right\}=\frac{pe^{t}}{1-qe^{t}}\qquad \text{desde que } \left|qe^{t}\right|<1$$

A primeira derivada da função geradora de momentos é:

$$\frac{dM(t)}{dt}=\frac{pe^{t}\left(1-qe^{t}\right)-\left(-qe^{t}\right)pe^{t}}{\left(1-qe^{t}\right)^{2}}$$

$$E(X)=\left.\frac{dM(t)}{dt}\right|_{t=0}=\frac{p(1-q)-(-q)p}{(1-q)^{2}}=\frac{p}{p^{2}}=\frac{1}{p}$$

Problema 5 (20 pontos)

No estacionamento de uma certa Universidade, observou-se que a proporção de carros de diversas marcas é:
Fiat = 18%
Volkswagen = 15%
Ford = 12%
GM = 10%
Outros (brasileiros) = 35%
Outros (importados) = 10%

Existem 30 carros na fila do estacionamento. Calcule as seguintes probabilidades:

a) De encontrar exatamente 5 carros da Fiat, 5 da Volkswagen, 4 da Ford, 3 da GM e 10 de outras marcas brasileiras.
b) De encontrar exatamente 5 carros da Fiat, 5 da Volkswagen, 4 da Ford e 3 da GM.
c) De encontrar exatamente 5 carros da Fiat, 5 da Volkswagen e 4 da Ford.
d) De encontrar exatamente 5 carros da Fiat.

Solução

Sejam:
X_1, X_2 , X_3, X_4, X_5 e X_6 o número de carros Fiat, VW, Ford, GM, “outros nacionais” e “outros importados” na fila do estacionamento.

a)
$$\Pr\left(X_1=5,X_2=5,X_3=4,X_4=3,X_5=10,X_6=3\right)=$$
$$=\frac{30!}{(5!)(5!)(4!)(3!)(10!)(3!)}(0.18)^{5}(0.15)^{5}(0.12)^{4}(0.10)^{3}(0.35)^{10}(0.10)^{3}=0.000482$$

b)
$$\Pr\left(X_1=5,X_2=5,X_3=4,X_4=3,\text{outros}=13\right)=$$
$$=\frac{30!}{(5!)(5!)(4!)(3!)(13!)}(0.18)^{5}(0.15)^{5}(0.12)^{4}(0.10)^{3}(0.45)^{13}=0.001897$$

c)
$$\Pr\left(X_1 = 5, X_2 = 5, X_3 = 4, \text{outros} = 16\right) =$$
$$= \frac{30!}{(5!)(5!)(4!)(16!)}(0.18)^5(0.15)^5(0.12)^4(0.55)^{16} = 0.007653$$

d)
$$\Pr\left(X_1 = 5, \text{outros} = 25\right) =$$
$$= \frac{30!}{(5!)(25!)}(0.18)^5(0.82)^{25} = 0.18860$$

Tabela – Função de Distribuição N(0,1)

z	Φ(z)		z	Φ (z)		z	Φ (z)
0,0000	50,00%		1,0000	84,13%		2,0125	97,79%
0,0200	50.80%		1,0100	84.38%		2.0200	97.83%
0,0300	51.20%		1.0167	84.54%		2.0300	97.88%
0,0400	51.60%		1.0250	84.73%		2.0400	97.93%
0,0500	51,99%		1,0500	85,31%		2.0412	97.94%
0,1000	53,98%		1,0553	85,44%		2.0500	97.98%
0,1500	55,96%		1,1000	86,43%		2,1000	98,21%
0,2000	57,93%		1,1180	86,82%		2,2000	98,61%
0,2236	58,85%		1,1475	87,44%		2,2361	98,73%
0,2500	59,87%		1,1500	87,49%		2,3000	98,93%
0,3000	61,79%		1,1553	87,60%		2,3263	99,00%
0,3015	61,85%		1,2000	88,49%		2,3333	99,02%
0,3475	63,59%		1,2060	88,61%		2,4000	99,18%
0,3492	63,65%		1,2200	88,88%		2,5000	99,38%
0,3500	63,68%		1,2500	89,44%		2,5500	99,46%
0,4000	65,54%		1.2700	89.79%		2,5628	99,48%
0.4167	66.16%		1,2816	90,00%		2,6000	99,53%
0,4307	66,67%		1,3000	90,32%		2,6500	99,60%
0,4500	67,36%		1,3333	90,88%		2,6667	99,62%
0,5000	69,15%		1,3750	91.54%		2,6833	99,64%
0,5500	70,88%		1,4000	91,92%		2,7000	99,65%
0,5774	71.81%		1,4468	92,60%		2.7500	99.70%
0,6000	72,57%		1,4500	92,65%		2,8000	99,74%
0,6250	73,40%		1,5000	93,32%		2,9000	99,81%
0,6500	74,22%		1,5500	93,94%		2,9500	99,84%
0,6667	74,75%		1,5811	94,31%		3,0000	99,87%
0,6708	74,88%		1,6000	94,52%		3,1000	99,90%
0,7000	75,80%		1,6450	95,00%		3,1500	99,92%
0,7500	77,34%		1,6667	95,22%		3,2000	99,93%
0,8000	78,81%		1,7000	95,54%			
0,8333	79.77%		1,8000	96,41%			
0,8500	80,23%		1.8500	96.78%			
0.8666	80.69%		1,9000	97,13%			
0,8944	81,45%		1.9500	97.44%			
0,9000	81,59%		1,9600	97,50%			
0,9167	82,03%		1,9800	97,61%			
0,9500	82,89%		2,0000	97,72%			
0,9500	82,89%		2,0100	97,78%			
0,9750	83,52%						
0,9800	83,65%						
0,9900	83,89%						