IND 1115 – Inferência Estatística – Semestre 2008.01
Teste 4 – 03/07/2008
Nome: GABARITO

NOTA: O FORMULÁRIO ESTÁ NO FINAL DA PROVA

PROBLEMA 1 (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

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

PROBLEMA 2 (25 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 consistente? É não tendencioso?

Solução

a) A verossimilhança é:

$$L(p)=\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i} = p^{n\overline{X}}(1-p)^{n-n\overline{X}}$$

A log-verossimilhança é:

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

Derivando a log-verossimilhança com relação a p leva a:

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

$$\Rightarrow \hat{p}=\overline{X}$$

é o estimador de máxima verossimilhança de p.

$\sum X_i = n\overline{X}$ tem distribuição Binomial(n,p) e portanto sua média é n.p e sua variância é npq.

Logo, a média e variância de $\overline{X}$ são, respectivamente, p e pq/n.

b) O MLE é consistente? É não tendencioso?

O estimador de máxima verossimilhança é não tendencioso para p, como mencionado acima.

O erro quadrático médio deste estimador é apenas a sua variância, que é p.q/n.

Como a variância tende a zero quando n tende a infinito, o MLE é consistente.

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 (35 pontos)

O consumo mensal em minutos por conta de celular numa certa região é uma v.a. Normal com média 40 minutos e desvio padrão 12 minutos.

a) Qual a probabilidade de alguém usar o celular menos de 50 minutos?

b) Qual a probabilidade de alguém usar o celular mais de 35 minutos?

c) Quantos minutos por mês alguém deve passar no celular para estar entre os 10% que mais usam o aparelho?

d) Quantos minutos por mês alguém deve passar no celular para estar entre os 5% que MENOS usam o aparelho?

Toma-se uma amostra de 16 usuários de celular.

e) Qual a probabilidade do tempo médio de uso na amostra exceder 45 minutos?

f) Qual a probabilidade do maior tempo de uso na amostra ser menor que 50 minutos?

g) Qual a probabilidade do menor tempo de uso na amostra ser menor que 24 minutos?

Solução

X = consumo em minutos ~ $N(40, 12^2)$

a) Pr(X < 50) = $\Pr\left(\frac{X-40}{12} < \frac{50-40}{12}\right) = \Pr(Z < 10/12) = \Phi\left(\frac{5}{6}\right) = \Phi(0.8333) = 0.7977$

b) Pr(X > 35) = $\Pr\left(\frac{X-40}{12} > \frac{35-40}{12}\right) = 1 - \Pr(Z > -5/12) = 1 - \Phi\left(\frac{-5}{12}\right) =$

$= 1 - \Phi(-0.4167) = \Phi(+0.4167) = 0.6615$

c) Para ser um dos 10% que mais usam o aparelho, a variável normalizada é z = 1.2816

Logo:

$$\frac{X-40}{12} = 1.2816 \Leftrightarrow X = 40 + 12(1.2816) = 55.38 \text{ minutos}$$

d) Para estar entre os 5% que menos usam o celular, a variável normalizada é z = -1.645 e então:

$$\frac{X-40}{12} = -1.645 \Leftrightarrow X = 40 + 12(-1.645) = 20.26 \text{ minutos}$$

Agora considere uma amostra de 16 usuários de celular. O tempo médio é uma variável Normal com média 40 minutos e variância $(12)^2/16 = 9$.

e) $\Pr(\overline{X} > 45) = \Pr\left(\frac{\overline{X}-40}{\sqrt{9}} > \frac{45-40}{\sqrt{9}}\right) = \Pr\left(Z > \frac{45-40}{3}\right) = \Pr(Z > 1.6667) =$

$= 1 - \Phi(1.6667) = 1 - 0.9522 = 0.0478$

f) Seja V = Max(X_1, X_2, ..., X_{16}). Então Pr(V < 50) = Pr(X_1 < 50, X_2 < 50, ..., X_{16} < 50) e como os X_i´s são iid, esta probabilidade é igual a $\{\Pr(X_1 < 50)\}^{16}$. Mas:

$$\Pr(X_1 < 50) = \Pr\left(\frac{X_1-40}{12} < \frac{50-40}{12}\right) = \Pr\left(Z < \frac{50-40}{12}\right) = \Phi(0.8333) = 0.7977$$

(vide item a))

E então: Pr(V < 50) = $(0.7977)^{16}$ = 0.0269

g) Seja U = Min(X_1, X_2, ..., X_{16}). Então:

Pr(U < 24) = 1 - Pr(U ≥ 24) = 1- Pr(X_1 ≥ 24, X_2 ≥ 24, ..., X_{16} ≥ 24) = 1 – $\{\Pr(X_1 \geq 24)\}^{16}$ =

$$= 1 - \left\{\Pr\left(\frac{X_1-40}{12} > \frac{24-40}{12}\right)\right\}^{16} = 1 - \{1 - \Phi(-1.3333)\}^{16} = 1 - \{\Phi(+1.3333)\}^{16} = 1 - (0.9088)^{16} = 0.7835$$

Tabela – Função de Distribuição N(0,1)

z	Φ(z)		z	Φ (z)		z	Φ (z)
0,0000	50,00%		0,9650	83,27%		1,9500	97,44%
0,0100	50,40%		0,9700	83,40%		1,9600	97,50%
0,0200	50,80%		0,9750	83,52%		1,9800	97,61%
0,0300	51,20%		0,9800	83,65%		1,9900	97,67%
0,0500	51,99%		0,9882	83,85%		2,0000	97,72%
0,1000	53,98%		0,9900	83,89%		2,0100	97,78%
0,1042	54,15%		1,0000	84,13%		2,0125	97,79%
0,1500	55,96%		1,0100	84,38%		2,0200	97,83%
0,2000	57,93%		1,0167	84,54%		2,0300	97,88%
0,2236	58,85%		1,0250	84,73%		2,0400	97,93%
0,2500	59,87%		1,0300	84,85%		2,0412	97,94%
0,2887	61,36%		1,0500	85,31%		2,0500	97,98%
0,3000	61,79%		1,0553	85,44%		2,1000	98,21%
0,3015	61,85%		1,0607	85,56%		2,1875	98,56%
0,3333	63,06%		1,1000	86,43%		2,2000	98,61%
0,3475	63,59%		1,1500	87,49%		2,2361	98,73%
0,3492	63,65%		1,1553	87,60%		2,3000	98,93%
0,3500	63,68%		1,1667	87,83%		2,3263	99,00%
0,4000	65,54%		1,2000	88,49%		2,3333	99,02%
0,4167	66,16%		1,2200	88,88%		2,4000	99,18%
0,4307	66,67%		1,2500	89,44%		2,5000	99,38%
0,4500	67,36%		1,2700	89,79%		2,5500	99,46%
0,5000	69,15%		1,2816	90,00%		2,5628	99,48%
0,5500	70,88%		1,3000	90,32%		2,6000	99,53%
0,5774	71,81%		1,3333	90,88%		2,6500	99,60%
0,6000	72,57%		1,3750	91,54%		2,6667	99,62%
0,6250	73,40%		1,4000	91,92%		2,6833	99,64%
0,6500	74,22%		1,4434	92,55%		2,7000	99,65%
0,6667	74,75%		1,4468	92,60%		2,7500	99,70%
0,7000	75,80%		1,4500	92,65%		2,8000	99,74%
0,7071	76,03%		1,5000	93,32%		2,9000	99,81%
0,7500	77,34%		1,5500	93,94%		2,9500	99,84%
0,8000	78,81%		1,6000	94,52%		3,0000	99,87%
0,8333	79,77%		1,6450	95,00%		3,1000	99,90%
0,8400	79,95%		1,6667	95,22%		3,1500	99,92%
0,8500	80,23%		1,7000	95,54%		3,1667	99,92%
0,8660	80,68%		1,7917	96,34%		3,2000	99,93%
0,9000	81,59%		1,8000	96,41%		3,8333	99,99%
0,9332	82,46%		1,8333	96,66%			
0,9500	82,89%		1,8500	96,78%			
0,9600	83,15%		1,9000	97,13%			

IND 1115 - Inferência Estatística
Profa. Mônica Barros
FORMULÁRIO P4

Nome	Densidade ou Função de Probabilidade	Média	Variância	fgm
Uniforme	$f(x)=\dfrac{1}{b-a} \quad \text{se } a \le x \le b$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	Não é útil
Exponencial	$f(x)=\lambda.\exp(-\lambda.x) \quad \text{onde } \lambda>0 \text{ e } x \ge 0$	$1/\lambda$	$1/\lambda^2$	$\left(\dfrac{\lambda}{\lambda-t}\right)$ se $t<\lambda$
Gama	$f(x)=\begin{cases}\dfrac{\beta^\alpha . x^{\alpha-1}}{\Gamma(\alpha)}.e^{-\beta x}, & \text{onde x} \ge 0 \\ 0 & \text{se x} < 0\end{cases}$	$\dfrac{\alpha}{\beta}$	$\dfrac{\alpha}{\beta^2}$	$\left(\frac{\beta}{\beta-t}\right)^\alpha$ se $t<\beta$
Qui-Quadrado	$f(x)=\dfrac{1}{2^{n/2}.\Gamma\left(\frac{n}{2}\right)}.x^{\frac{n}{2}-1}.e^{-\frac{x}{2}}$	n	2n	$\left(\dfrac{1}{1-2t}\right)^{n/2}$ se $t<1/2$
Normal	$f(x)=\dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{\dfrac{-(x-\mu)^2}{2\sigma^2}\right\}$	μ	σ^2	$\exp\left\{\mu.t+\dfrac{\sigma^2 t^2}{2}\right\}$
Binomial	$f(x)=\Pr(X=x)=\binom{n}{x}p^x(1-p)^{n-x}=\dfrac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$ para $x=0,1,2,\ldots,n$	n.p	n.p.q	$(pe^t+q)^n$
Hipergeométrica	$f(x)=\Pr(X=x)=\dfrac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$	$n.\left(\dfrac{r}{N}\right)$	$n.\left(\dfrac{r}{N}\right)\left(1-\dfrac{r}{N}\right)\left(\dfrac{N-n}{N-1}\right)$	Não é útil
Geométrica	$f(n)=\Pr(X=n)=(1-p)^{n-1}p$ onde n = 1, 2, 3, ...	1/p	q/p^2	$M(t)=\dfrac{pe^t}{1-qe^t}$
Poisson	$\Pr(X=x)=f(x)=\dfrac{\lambda^x.e^{-\lambda}}{x!}$ onde x = 0, 1, 2, ...	λ	λ	$E(e^{tX})=e^{\lambda.(e^t-1)}$
Binomial Negativa	$f(x)=\Pr(X=x)=\binom{x-1}{r-1}.p^r.q^{x-r}$	r/p	$r.q/p^2$	$M(t)=\left(\dfrac{pe^t}{1-qe^t}\right)^r$

onde x = r, r + 1, r + 2, ...

Resultados Matemáticos

Série Geométrica

$$\sum_{k=0}^{\infty} a^k = 1+a+a^2+a^3+\ldots\ldots=\frac{1}{1-a} \quad \text{desde que } |a|<1$$

Teorema Binomial

$$(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \quad \text{onde a e b são número reais e k, n são inteiros} \ge 0$$

Série de Taylor da Exponencial

$$e^x=\sum_{k=0}^{\infty}\left(\frac{x^k}{k!}\right)=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots+\frac{x^k}{k!}+\ldots\ldots$$

Função de Distribuição (F(x))

1) $F(x) = \Pr(X \le x)$
2) $0 \le F(x) \le 1$
3) F(x) é uma função não decrescente
4) Lim F(x) = 1 se $x \to +\infty$
5) Lim F(x) = 0 se $x \to -\infty$
6) Se X é uma v.a. contínua, F(x) é contínua. Se X é discreta, F(x) é descontínua

$$f(x) = \frac{dF(x)}{dx}$$

Relação entre densidade e função de distribuição

Definição: k-ésimo momento

$$E(X^k) = \begin{cases} \int_{-\infty}^{\infty} x^k . f(x) dx & \text{se X é v.a. contínua} \\ \sum_{\text{todo x}} x^k . f(x) = \sum_{\text{todo x}} x^k . \Pr(X = x) & \text{se X é v.a. discreta} \end{cases}$$

Definição: k-ésimo momento central

$$E((X-\mu)^k) = \begin{cases} \int_{-\infty}^{\infty} (x-\mu)^k . f(x) dx & \text{se X é v.a. contínua} \\ \sum_{\text{todo x}} (x-\mu)^k . f(x) = \sum_{\text{todo x}} (x-\mu)^k . \Pr(X = x) & \text{se X é v.a. discreta} \end{cases}$$

Em particular, se k = 1: E(X – m) = 0, ou seja, o primeiro momento central é sempre nulo.

Definição: Média ou Valor Esperado de X

$$\mu = E(X) = \begin{cases} \int_{-\infty}^{\infty} x . f(x) dx & \text{se X é v.a. contínua} \\ \sum_{\text{todo x}} x . f(x) = \sum_{\text{todo x}} x . \Pr(X = x) & \text{se X é v.a. discreta} \end{cases}$$

Definição: Variância

$$\sigma^2 = VAR(X) = E\left((X-\mu)^2\right) = \begin{cases} \int_{-\infty}^{\infty} (x-\mu)^2 . f(x) dx & \text{se X contínua} \\ \sum_{\text{todo x}} (x-\mu)^2 . f(x) = \sum_{\text{todo x}} (x-\mu)^2 . \Pr(X = x) & \text{se X discreta} \end{cases}$$

Fórmula alternativa para o cálculo da variância

$$\sigma^2 = VAR(X) = E(X^2) - \mu^2$$

Definição: Desvio padrão

$$\sigma = \sqrt{\sigma^2} = \sqrt{VAR(X)}$$

Definição: Valor esperado de uma função de uma variável aleatória

$$E(u(X)) = \begin{cases} \int_{-\infty}^{\infty} u(x) . f(x) dx & \text{se X é v.a. contínua} \\ \sum_{\text{todo x}} u(x) . f(x) = \sum_{\text{todo x}} u(x) . \Pr(X = x) & \text{se X é v.a. discreta} \end{cases}$$

Propriedade – linearidade do valor esperado:

E{a.u(X) + b.v(X)} = a E {u(X)} + b E {v(X)}

Propriedades – Média e Variância de constantes e funções lineares

Sejam a e b constantes, e X uma variável aleatória qualquer. Então:

1-) E(a.X + b) = a.E(X) + b
2-) E(a) = a
3-) VAR(a.X+ b) = a^2.VAR(X)
4-) VAR(a) = 0

Poisson como aproximação da Binomial

Se X é Binomial(n, p), onde n é grande e p é pequeno (n > 20 e n.p < 5), pode-se aproximar as probabilidades Binomiais por probabilidades Poisson usando uma Poisson com a mesma média, isto é, usando λ = n.p.

Função Geradora de Momentos (fgm)

$$M(t) = E(e^{tX}) = \begin{cases} \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{se X é v.a. contínua} \\ \sum_{\text{todo x}} e^{tx} . f(x) = \sum_{\text{todo x}} e^{tx} . \Pr(X = x) & \text{se X é v.a. discreta} \end{cases}$$

Relação entre Momentos e fgm

$$M^{(k)}(0) = \left. \frac{d^k M(t)}{dt^k} \right|_{t=0} = E(X^k)$$

Fórmula da Convolução

Seja Y= X_1+X_2 onde X_1 e X_2 são variáveis independentes com densidade conjunta $f(x_1, x_2) = f_1(x_1).f_2(x_2)$. A densidade (ou função de probabilidade de Y) é dada por:

a) No caso contínuo

$$g(y) = \int_{-\infty}^{\infty} f_1(x_1).f_2(y-x_1)dx_1 = \int_{-\infty}^{\infty} f_1(y-x_2).f_2(x_2)dx_2$$

b) No caso discreto

$$g(y) = \sum_{\text{todo } x_1} f_1(x_1).f_2(y-x_1) = \sum_{\text{todo } x_2} f_1(y-x_2).f_2(x_2)$$

Combinações Lineares de variáveis INDEPENDENTES

Sejam $X_1, X_2, \ldots, X_n$ **independentes** com médias $\mu_1, \mu_2, \ldots, \mu_n$ e variâncias $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$.

Seja: $Y = a_0 + \sum_{i=1}^{n} a_i X_i$

Então, a **média de Y** é:

$$E(Y) = E\left(a_0 + \sum_{i=1}^{n} a_i.X_i\right) = a_0 + \sum_{i=1}^{n} a_i.E(X_i) =$$

$$= a_0 + \sum_{i=1}^{n} a_i.\mu_i \quad \text{onde } \mu_i = E(X_i)$$

A **variância de Y** é:

$$VAR(Y) = \sum_{i=1}^{n} a_i^2.VAR(X_i) = \sum_{i=1}^{n} a_i^2.\sigma_i^2$$

$$\text{onde } VAR(X_i) = \sigma_i^2$$

E a **fgm de Y** é:

$$M_Y(t) = E\left(e^{tY}\right) = E\left(e^{t(a_0 + a_1X_1 + \ldots + a_nX_n)}\right)$$

$$= E\left(e^{ta_0} e^{ta_1X_1} \ldots . e^{ta_nX_n}\right) =$$

$$= e^{ta_0} E\left(e^{ta_1X_1} \ldots . e^{ta_nX_n}\right)$$

e como conseqüência da independência

$$= e^{ta_0} E\left(e^{ta_1X_1}\right) E\left(e^{ta_2X_2}\right) \ldots E\left(e^{ta_nX_n}\right) =$$

$$e^{ta_0} M_{X_1}(ta_1) M_{X_2}(ta_2) \ldots M_{X_n}(ta_n)$$

Combinações Lineares de variáveis DEPENDENTES

Suponha que as médias e variâncias dos X_i´s são como no caso anterior, mas agora eles são DEPENDENTES, de tal forma que: $COV(X_i, X_j) = COV(X_j, X_i) = \rho_{ij}.\sigma_i.\sigma_j$ onde ρ_{ij} é o coeficiente de correlação entre X_i e X_j.

Seja Y definido como acima. $Y = a_0 + \sum_{i=1}^{n} a_i X_i$

Então E(Y) é o mesmo que no caso de variáveis dependentes, MAS:

$$VAR(Y) = \sum_{i=1}^{n} a_i^2.VAR(X_i) + \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j COV\left(X_i, X_j\right)$$

$$= \sum_{i=1}^{n} a_i^2.\sigma_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \rho_{ij} \sigma_i \sigma_j \quad \text{onde no 2o. termo } i \neq j$$

onde ρ_{ij} é o coeficiente de correlação entre X_i e X_j

Portfolio

Combinação linear de ativos onde soma dos pesos = 1. Nos nossos exemplos estamos supondo que todos os pesos são positivos.

A média e a variância do portfolio podem ser obtidas diretamente das expressões acima para combinações lineares de variáveis dependentes. NOTAR que, acima no termo da covariância aparecem $a_i.a_j$ e $a_j.a_i$.

Função Gama

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1}.e^{-t} dt$$

Propriedades da Função Gama

1) $\Gamma(n) = (n-1).\Gamma(n-1)$ para $n > 1$
2) $\Gamma(n) = (n-1)!$ se n é inteiro > 1
3) $\Gamma(1) = 0! = 1$
4) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Padronização de uma variável aleatória

Se X tem média a e variância b^2 então $Z = (X-a)/b$ tem média 0 e variância 1. Se, além disso, X é Normal, Z também é Normal.

Cálculo de probabilidades para variáveis Normais

Se X é uma variável Normal com média μ e desvio padrão σ então:

$$\Pr(a \le X \le b) = \Pr\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) = \Pr\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Simetria da tabela da N(0,1)

$\Phi(-z) = 1 - \Phi(z)$ se $z > 0$

Combinações Lineares de Variáveis Normais

Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias **independentes**, onde $X_i \sim N(\mu_i, \sigma_i^2)$ e seja $Y = X_1 + X_2 + \ldots + X_n$.

Então Y tem distribuição Normal com média μ_y e variância σ_y^2 dadas por:

$$\mu_y = \sum_{i=1}^{n} \mu_i \quad \text{e} \quad \sigma_y^2 = \sum_{i=1}^{n} \sigma_i^2$$

Casos particulares: se todos os X_i´s acima são iid com média μ e variância σ^2 então a soma é Normal com média $n.\mu$ e variância $n.\sigma^2$ e a média amostral é Normal com média μ e variância σ^2/n.

Densidade Lognormal

Se X é N(μ, σ^2) então Y = e^X é Lognormal. Pode-se provar que E(Y) = exp(μ + σ^2/2) e

$$VAR(Y) = \exp\left(2\mu + \sigma^2\right).\left(e^{\sigma^2} - 1\right)$$

Densidade Beta

$$f(x) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} x^{m-1}(1-x)^{n-1} \quad \text{onde } 0 < x < 1 \text{ e m, n inteiros} \geq 1$$

Se X ~ Beta (m, n) então:

$$E(X) = \frac{m}{m+n} \qquad VAR(X) = \frac{mn}{(m+n+1)\ (m+n)^2} \qquad \text{e} \qquad E(X^k) = \frac{\Gamma(k+m)\Gamma(m+n)}{\Gamma(m)\Gamma(k+m+n)}$$

Teorema

Sejam X_1, X_2, …., X_n variáveis aleatórias independentes com densidade Unif(0,1). Seja Y_r o r-ésimo número (em ordem crescente) dentre os valores observados de X_1, X_2, …., X_n , de tal forma que Y_1 é o mínimo, Y_2 é o segundo menor, ..., Y_n é o máximo da amostra. Então Yr tem densidade Beta com parâmetros r e n – r + 1.

Teorema Central do Limite (TCL)

Sejam X_1, X_2, …., X_n variáveis aleatórias **independentes com densidades quaisquer**, onde E(X_i) = μ_i e VAR(X_i) = σ_i^2 ambas finitas.
Seja Y = X_1 + X_2 + … + X_n . Então Y, devidamente padronizado para ter média zero e variância 1 é aproximadamente Normal, desde que o número de termos na soma seja suficientemente grande. O teorema é importante porque serve para X_i´s com QUALQUER distribuição, contínua ou discreta.
Casos Particulares do TCL – se os X´s são iid, todos têm mesma média e variância.

Teorema de DeMoivre e Laplace

Caso particular do TCL aplicado à Binomial. Historicamente apareceu muito antes do TCL. Usa-se quando uma Binomial tem n grande e p proóximo de ½. Se uma Binomial tem n grande e p pequeno, é melhor usar a aproximação Poisson ao invés de DeMoivre e Laplace. Um problema na aplicação de DeMoivre e Laplace é o fato de estarmos aproximando uma variável discreta (Binomial) por uma contínua (Normal). Daí surge a necessidade da correção de continuidade, do contrário não teríamos como calcular Pr(Y=k), por exemplo, usando a aproximação.
Seja Y ~ Bin(n, p) onde n é "grande" e p não está próximo de zero. Então:

$$Z = \frac{Y - E(Y)}{\sqrt{VAR(Y)}} = \frac{Y - np}{\sqrt{np(1-p)}}$$ **tem aproximadamente uma distribuição N(0,1).**

Correção de Continuidade

Quantidade desejada na distribuição Binomial	Quantidade Calculada através da correção de continuidade	Expressão aproximada usando a densidade Normal
Pr(Y = y)	Pr(y - 0.5 ≤ Y ≤ y + 0.5)	$\Phi\left(\frac{y+0.5-np}{\sqrt{npq}}\right) - \Phi\left(\frac{y-0.5-np}{\sqrt{npq}}\right)$
Pr(Y ≤ y)	Pr(Y ≤ y +0.5)	$\Phi\left(\frac{y+0.5-np}{\sqrt{npq}}\right)$
Pr(Y < y) = Pr(Y ≤ y-1)	Pr(Y ≤ y - 1 + 0.5)	$\Phi\left(\frac{y-1+0.5-np}{\sqrt{npq}}\right) = \Phi\left(\frac{y-0.5-np}{\sqrt{npq}}\right)$
Pr(Y ≥ y)	Pr(Y ≥ y - 0.5)	$1-\Phi\left(\frac{y-0.5-np}{\sqrt{npq}}\right)$
Pr(Y > y) = Pr(Y ≥ y +1)	Pr(Y ≥ y + 1 - 0.5)	$1-\Phi\left(\frac{y+1-0.5-np}{\sqrt{npq}}\right)$
Pr(a ≤ Y ≤ b)	Pr(a - 0.5 ≤ Y ≤ b + 0.5)	$\Phi\left(\frac{b+0.5-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-0.5-np}{\sqrt{npq}}\right)$

Teorema – Aditividade da Qui-quadrado

Sejam X_1, X_2, …., X_n variáveis aleatórias **independentes** e X_i é Qui-Quadrado com k_i graus de liberdade. Então Y = X_1 + X_2 + … + X_n é também Qui-Quadrado, com k_1 + k_2 + …+ k_n graus de liberdade.

Teorema – Relação entre Normal e Qui-Quadrado

Seja Z ~ N(0,1) . Então V = Z^2 tem densidade Qui-quadrado com 1 grau de liberdade.

Definição – Amostra Aleatória

É um conjunto de observações independentes e identicamente distribuídas.

Teorema – Distribuição da Média e da Variância amostrais numa amostra Normal

Seja X_1, X_2, …., X_n uma amostra da densidade N(μ, σ^2).

Sejam $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ a média amostral e $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ a variância amostral

Então $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ e $\frac{\sum \left(X_i - \bar{X}\right)^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ e $\bar{X}$ e S^2 são independentes

Deste resultado deduz-se que:

$E(S^2) = \sigma^2$

$VAR(S^2) = \frac{2\sigma^4}{n-1}$

A densidade t de Student – definição

Uma variável t com k graus de liberdade é obtida através de:

$T = \frac{Z}{\sqrt{V/k}}$ onde Z e V são independentes, Z é N(0,1) e V é Qui-Quadrado com k graus de liberdade.

À medida que os graus de liberade da distribuição t crescem, ela se aproxima de uma N(0,1).

A distribuição t e amostras Normais

Seja $X_1, X_2, \ldots, X_n$ uma amostra da densidade $N(\mu, \sigma^2)$. Considere e média e variância amostrais como já definidas. Então:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{\sqrt{n}\left(\bar{X} - \mu\right)}{S} \sim t_{n-1}$$

Estatística

Estimação Pontual – encontrar “chutes” (estimadores) para parâmetros desconhecidos.
Principais Métodos de Estimação
Método dos momentos
Método de máxima verossimilhança
Método dos mínimos quadrados

Método dos Momentos – a idéia é igualar os momentos amostrais aos momentos da distribuição ($E(X^k)$) tantas vezes quanto necessário até encontrar uma solução única para todos os parâmetros desconhecidos. Se apenas um parâmetro é desconhecido, basta fazer isso uma vez.

Função de Verossimilhança ($L(\theta)$)

A função de verossimilhança é a densidade conjunta encarada como função do parâmetro θ. Isto é: $L(\theta) = f\left(x_1, x_2, \ldots, x_n\right) = \prod_{i=1}^{n} f\left(x_i, \theta\right)$

Log-verossimilhança ($l(\theta)$)

É o logaritmo na base e da verossimilhança.

Método da Máxima Verossimilhança

Consiste em achar um estimador que maximiza a verossimilhança (ou, de modo equivalente, a log-verossimilhança). Em geral, ele é encontrado por Cálculo, resolvendo-se a equação $dl/d\theta = 0$, mas existem exceções, como a densidade Uniforme.

Definição (Estimador não tendencioso)

Seja T um estimador para o parâmetro θ de uma densidade $f(x, \theta)$. T é chamado de *não tendencioso* se $E(T) = \theta$, do contrário T é dito tendencioso.

Definição (Erro Quadrático Médio)

O erro quadrático médio do estimador T é definido como: $MSE(T) = E\{ (T - \theta)^2 \} = VAR(T) + \{ BIAS(T)\}^2$ onde θ é o parâmetro que T pretende estimar, VAR(T) é a variância de T e BIAS(T) é a tendência ou viés de T, definido como $BIAS(T) = E(T) - \theta$.

Definição (Estimador consistente)

Um estimador T é consistente se seu MSE vai para zero quando n vai para infinito, onde n é o tamanho da amostra.

Método dos Momentos

Seja $E(X^k)$ o k-ésimo momento da distribuição (k =1,2,...).
Seja:

$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ o k-ésimo momento amostral

Faça $E(X^k) = M_k$ para k = 1,2,....
Faça isto para quantos k's forem necessários até obter soluções únicas para os parâmetros desconhecidos