

IND 1115 – Inferência Estatística – Semestre 2008.01

Teste 3 – 26/06/2008

Nome: GABARITO

NOTA: O FORMULÁRIO ESTÁ NO FINAL DA PROVA

FOLHA DE RESPOSTAS – PREENCHA A CANETA E NÃO RASURE!!!

Nos problemas de múltipla escolha, marque com um X na folha de respostas a alternativa que você considera correta. O formulário está no final da prova.

PROBLEMA 1

	A	B	C	D
Item 1		<input checked="" type="checkbox"/>		
Item 2			<input checked="" type="checkbox"/>	

PROBLEMA 3

	A	B	C	D
Item 2				<input checked="" type="checkbox"/>

PROBLEMA 1 (25 pontos)

Pelo Teorema Central do Limite:

1) Se X é Poisson(36) então $\Pr(X < 45)$ é, aproximadamente:

	A	B	C	D
Item 1	0.0668	0.9332	0.5987	0.8556

2) Se X é Qui-Quadrado com 50 graus de liberdade então: $\Pr(X > 45)$ é aproximadamente:

	A	B	C	D
Item 2	0.3085	0.7603	0.6915	0.2387

Solução

1) X é Poisson(36), então pode ser pensada como a soma de 36 Poisson(1), todas independentes, o que justifica o uso do Teorema Central do Limite. Note que $E(X) = \text{VAR}(X) = 36$ neste caso. Assim:

$$Z = \frac{X - 36}{\sqrt{36}} = \frac{X - 36}{6} \text{ é aproximadamente } N(0,1).$$

$$\text{Logo, } \Pr(X < 45) = \Pr\left(\frac{X - 36}{6} < \frac{45 - 36}{6}\right) \approx \Pr\left(Z < \frac{9}{6}\right) = \Phi\left(\frac{9}{6}\right) = \Phi(1.5) = 0.9332$$

2) Agora X é Qui-Quadrado com 50 graus de liberdade, e pode ser pensada como a soma de 50 variáveis iid Qui-quadrado com 1 grau de liberdade, o que justifica o uso do Teorema Central do Limite. Note que $E(X) = 50$ e $\text{VAR}(X) = 2(50) = 100$ neste caso. Assim:

$$Z = \frac{X - 50}{\sqrt{100}} = \frac{X - 50}{10} \text{ é aproximadamente } N(0,1).$$

$$\Pr(X > 45) = \Pr\left(\frac{X - 50}{10} > \frac{45 - 50}{10}\right) \approx \Pr\left(Z > \frac{-5}{10}\right) = 1 - \Phi(-0.5) = \Phi(+0.5) = 0.6915$$

PROBLEMA 2 (25 pontos)

Sejam X_1, X_2, \dots, X_n iid $N(0, \sigma^2)$.

a) Qual a média e variância de X_i^2 ?

b) Como aproximar $\Pr\{X_1^2 + X_2^2 + \dots + X_n^2 \leq x\}$ em termos de $\Phi(\cdot)$?

DICA – qual a densidade do quadrado de uma $N(0,1)$?

Solução

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

Logo, sua média e variância são:

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

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

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

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

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

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

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

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

PROBLEMA 3 (25 pontos)

Considere uma amostra aleatória de tamanho n da seguinte função de probabilidade:

$$\Pr(X = x) = f(x) = p^x (1 - p) \quad \text{para } x = 0, 1, 2, 3, \dots$$

- 1) Encontre o estimador de máxima verossimilhança de p .
- 2) O MLE de $\Pr(X > 2)$ quando $\bar{X} = 1.4$ é:

	A	B	C	D
Item 2	0.3403	0.8015	0.5833	0.1985

Solução

A verossimilhança é:

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

A log-verossimilhança é:

$$l(p) = \ln L(p) = \left(\sum x_i \right) \log(p) + n \log(1-p)$$

Derivando em relação a p e igualando a zero temos:

$$\frac{\partial l}{\partial p} = \frac{n\bar{X}}{p} - \frac{n}{1-p} = 0 \Leftrightarrow \frac{n\bar{X}}{p} = \frac{n}{1-p} \Leftrightarrow \frac{\bar{X}}{p} = \frac{1}{1-p} \Leftrightarrow p = \bar{X} - p\bar{X}$$

$$\Leftrightarrow p + p\bar{X} = \bar{X} \Leftrightarrow p(1 + \bar{X}) = \bar{X}$$

$$\Leftrightarrow \hat{p} = \frac{\bar{X}}{1 + \bar{X}}$$

b) O MLE de $\Pr(X > 2)$ pode ser calculado pelo princípio de invariância.

$$\begin{aligned} \Pr(X > 2) &= \Pr(X \geq 3) = 1 - \Pr(X = 0) - \Pr(X = 1) - \Pr(X = 2) = 1 - (1-p) - p(1-p) - p^2(1-p) = \\ &= 1 - (1-p) \{1 + p + p^2\} \end{aligned}$$

Pelo princípio da invariância basta substituir p por seu MLE, que é, neste caso:

$$\hat{p} = \frac{\bar{X}}{1 + \bar{X}} = \frac{1.4}{1 + 1.4} = \frac{1.4}{2.4} = \frac{14}{24} = \frac{7}{12} = 0.5833$$

Então, o MLE desejado é:

$$\begin{aligned} \Pr(X > 2) &= 1 - (1 - \hat{p}) \{1 + \hat{p} + \hat{p}^2\} = 1 - (1 - 0.5833) \{1 + 0.5833 + (0.5833)^2\} = \\ &= 1 - (0.4167) \{1.9236\} = 1 - 0.8015 = 0.1985 \end{aligned}$$

PROBLEMA 4 (25 pontos)

Sejam X_1, X_2, X_3 com função de probabilidade conjunta Multinomial com parâmetros n, p_1, p_2, p_3 onde p_1, p_2, p_3 são funções de θ dadas por $p_1 = \theta^2, p_2 = 2\theta(1-\theta)$ e $p_3 = (1-\theta)^2$. **Então a VEROSSIMILHANÇA é:**

$$f(x_1, x_2, x_3) = \frac{n!}{x_1! x_2! x_3!} \theta^{x_1} \{2\theta(1-\theta)\}^{x_2} \{(1-\theta)^2\}^{x_3} \quad \text{onde } \theta^2 + 2\theta(1-\theta) + (1-\theta)^2 = 1 \quad \text{e } x_1 + x_2 + x_3 = n,$$

$0 < \theta < 1$ e os x_i 's são inteiros não negativos sujeitos à restrição de somarem a n.

Ache o estimador de máxima verossimilhança de θ .

DICA:

Note que só existe um parâmetro desconhecido (θ). Não se esqueça de usar as restrições!

Solução

A log-verossimilhança é:

$$\begin{aligned}
 l(p_1, p_2) &= k + x_1 \ln(\theta^2) + x_2 \ln(2\theta(1-\theta)) + (n - x_1 - x_2) \ln\{(1-\theta)^2\} = \\
 &= k + 2x_1 \cdot \ln(\theta) + x_2 \cdot \ln(2) + x_2 \cdot \ln(\theta) + x_2 \cdot \ln(1-\theta) + 2(n - x_1 - x_2) \ln(1-\theta) = \\
 &= k_2 + 2x_1 \cdot \ln(\theta) + x_2 \cdot \ln(\theta) + x_2 \cdot \ln(1-\theta) + 2(n - x_1 - x_2) \ln(1-\theta) = \\
 &= k_2 + \{2x_1 + x_2\} \cdot \ln(\theta) + \{x_2 + 2(n - x_1 - x_2)\} \cdot \ln(1-\theta)
 \end{aligned}$$

Onde k_2 não envolve θ e portanto irá sumir quando derivarmos, então não precisamos nos preocupar com ele.

Derivando com relação a θ e igualando a zero:

$$\begin{aligned}
 \frac{dl}{d\theta} &= \frac{2x_1 + x_2}{\theta} + (-1) \left\{ \frac{x_2 + 2(n - x_1 - x_2)}{1-\theta} \right\} = 0 \\
 \Leftrightarrow \frac{2x_1 + x_2}{\theta} &= \frac{x_2 + 2(n - x_1 - x_2)}{1-\theta} \\
 \Leftrightarrow \theta \{x_2 + 2(n - x_1 - x_2)\} &= (1-\theta) \{2x_1 + x_2\} \\
 \Leftrightarrow \theta \{x_2 + 2(n - x_1 - x_2) + 2x_1 + x_2\} &= 2x_1 + x_2 \\
 \Leftrightarrow \theta \{+2(n - x_1 - x_2) + 2x_1 + 2x_2\} &= 2x_1 + x_2 \\
 \Leftrightarrow \theta \{2n\} &= 2x_1 + x_2 \\
 \Leftrightarrow \hat{\theta} &= \frac{2x_1 + x_2}{2n}
 \end{aligned}$$

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 P3

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

onde $x = r, r + 1, r + 2, \dots$

Resultados Matemáticos

Série Geométrica

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

Teorema Binomial

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

Série de Taylor da Exponencial

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

Função de Distribuição (F(x))

- 1) $F(x) = \Pr(X \leq x)$
- 2) $0 \leq F(x) \leq 1$
- 3) $F(x)$ é uma função não decrescente
- 4) $\lim_{x \rightarrow +\infty} F(x) = 1$ se $x \rightarrow +\infty$
- 5) $\lim_{x \rightarrow -\infty} F(x) = 0$ se $x \rightarrow -\infty$
- 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

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

Definição: k-ésimo momento central

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

Em particular, se $k = 1$: $E(X - \mu) = 0$, ou seja, o primeiro momento central é sempre nulo.

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

Definição: Média ou Valor Esperado

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

Definição:

Variância

de X

Fórmula alternativa para o cálculo da variância

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

Definição: Desvio padrão

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

Definição: Valor esperado de uma função de uma variável aleatória

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

Propriedade – linearidade do valor esperado:

$$E\{a \cdot u(X) + b \cdot v(X)\} = a E\{u(X)\} + b E\{v(X)\}$$

Propriedades – Média e Variância de constantes e funções lineares

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

- 1-) $E(a \cdot X + b) = a \cdot E(X) + b$
- 2-) $E(a) = a$
- 3-) $\text{VAR}(a \cdot X + b) = a^2 \cdot \text{VAR}(X)$
- 4-) $\text{VAR}(a) = 0$

Poisson como aproximação da Binomial

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

Função Geradora de Momentos (fgm)

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

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 $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$. A densidade (ou função de probabilidade de Y) é dada por:

a) No caso contínuo

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

b) No caso discreto

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

Combinações Lineares de variáveis INDEPENDENTES

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

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:

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

onde $VAR(X_i) = \sigma_i^2$

E a fgm de Y é:

$$M_Y(t) = E(e^{tY}) = E\left(e^{t(a_0 + a_1 X_1 + \dots + a_n X_n)}\right) = E\left(e^{ta_0} e^{ta_1 X_1} \dots e^{ta_n X_n}\right) = e^{ta_0} E\left(e^{ta_1 X_1} \dots e^{ta_n X_n}\right)$$

e como consequência da independência

$$= e^{ta_0} E\left(e^{ta_1 X_1}\right) E\left(e^{ta_2 X_2}\right) \dots E\left(e^{ta_n X_n}\right) = e^{ta_0} M_{X_1}(ta_1) M_{X_2}(ta_2) \dots M_{X_n}(ta_n)$$

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: $COV(X_i, X_j) = \rho_{ij} \sigma_i \sigma_j$ 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:

$$VAR(Y) = \sum_{i=1}^n a_i^2 VAR(X_i) + \sum_{i=1}^n \sum_{j=1}^n a_i a_j COV(X_i, X_j) = \sum_{i=1}^n a_i^2 \sigma_i^2 + \sum_{i=1}^n \sum_{j=1}^n a_i a_j \rho_{ij} \sigma_i \sigma_j \quad \text{onde no 2o. termo } i \neq j$$

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

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

Propriedades da Função Gama

- 1) $\Gamma(n) = (n-1) \cdot \Gamma(n-1)$ 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:

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

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

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 $n \cdot \mu$ e variância $n \cdot \sigma^2$ e a média amostral é Normal com média μ e variância σ^2/n .

Densidade Lognormal

Se $X \in N(\mu, \sigma^2)$ então $Y = e^X$ é Lognormal. Pode-se provar que $E(Y) = \exp(\mu + \sigma^2/2)$ e $VAR(Y) = \exp(2\mu + \sigma^2) \cdot (e^{\sigma^2} - 1)$

Densidade Beta

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

Se $X \sim \text{Beta}(m, n)$ então:

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

Teorema

Sejam X_1, X_2, \dots, X_n variáveis aleatórias independentes com densidade $\text{Unif}(0,1)$. Seja Y_r o r -ésimo número (em ordem crescente) dentre os valores observados de X_1, X_2, \dots, 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 Y_r tem densidade Beta com parâmetros r e $n - r + 1$.

Teorema Central do Limite (TCL)

Sejam X_1, X_2, \dots, X_n variáveis aleatórias independentes com densidades quaisquer, onde $E(X_i) = \mu_i$ e $VAR(X_i) = \sigma_i^2$ ambas finitas. Seja $Y = X_1 + X_2 + \dots + X_n$. 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 próximo de $1/2$. 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 $Y \sim \text{Bin}(n, p)$ onde n é "grande" e p não está próximo de zero. Então:

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

Correção de Continuidade

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

Teorema – Aditividade da Qui-quadrado

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

Teorema – Relação entre Normal e Qui-Quadrado

Seja $Z \sim N(0,1)$. 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, \dots, X_n uma amostra da densidade $N(\mu, \sigma^2)$.

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

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

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

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

A distribuição t e amostras Normais

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

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

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: $L(\theta) = f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i, \theta)$

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 ($k = 1, 2, \dots$).

Seja:

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

Faça $E(X^k) = M_k$ para $k = 1, 2, \dots$

Faça isto para quantos k 's forem necessários até obter soluções únicas para os parâmetros desconhecidos