

IND 1115 – Inferência Estatística – Semestre 2007.02

Teste 1 – 20/09/2007

Nome: _____

NOTA: ESCREVA AS RESPOSTAS COMO FRAÇÕES OU COM 4 CASAS DECIMAIS

NOTA 2: O FORMULÁRIO ESTÁ NO FINAL DA PROVA

Problema 1 (25 pontos)

Seja X uma variável aleatória com função de probabilidade Geométrica com probabilidade p , isto é:

$$\Pr(X = x) = (1 - p)^{x-1} p \quad \text{onde } x = 1, 2, \dots$$

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

Solução

A função geradora de momentos de X é:

$$\begin{aligned} E(e^{tx}) &= \sum_{i=1}^{\infty} e^{tx} (1-p)^{x-1} p = p(1-p)^{-1} \sum_{i=1}^{\infty} \left\{ e^{tx} (1-p)^x \right\} = p(q)^{-1} \sum_{i=1}^{\infty} \left\{ (qe^t)^x \right\} = \\ &= \frac{p}{q} \left\{ \frac{qe^t}{1-qe^t} \right\} = \frac{pe^t}{1-qe^t} \quad \text{desde que } |qe^t| < 1 \end{aligned}$$

A primeira derivada da função geradora de momentos é:

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

A segunda derivada da fgm é:

$$\frac{d^2M}{dt^2} = \frac{d}{dt} \left\{ \frac{pe^t}{(1-qe^t)^2} \right\} = \frac{pe^t(1-qe^t)^2 - 2(1-qe^t)(-qe^t)pe^t}{(1-qe^t)^4}$$

O segundo momento é esta derivada avaliada em zero:

$$E(X^2) = \frac{p(1-q)^2 - 2(1-q)(-q)p}{(1-q)^4} = \frac{p^3 + 2p^2q}{p^4} = \frac{p + 2q}{p^2} = \frac{1+q}{p^2}$$

A variância de X é então:

$$\text{VAR}(X) = E(X^2) - \{E(X)\}^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$

Problema 2 (25 pontos)

Sejam X_1 e X_2 variáveis aleatórias independentes, onde X_1 tem função de probabilidade Poisson com média λ_1 e X_2 tem função de probabilidade Poisson com média λ_2 . Mostre, usando o método que você preferir, que $Y = X_1 + X_2$ tem função de probabilidade Poisson com média $\lambda_1 + \lambda_2$.

Solução

A fgm de uma v.a. Poisson com média λ é $M(t) = \exp\{\lambda(e^t - 1)\}$

A fgm de $Y = X_1 + X_2$ é então:

$M_Y(t) = E(e^{tY}) = E(e^{tX_1} e^{tX_2})$ e pela independência de X_1 e X_2 temos:

$$M_Y(t) = E(e^{tX_1} e^{tX_2}) = E(e^{tX_1})E(e^{tX_2}) = \exp\{\lambda_1(e^t - 1)\} \exp\{\lambda_2(e^t - 1)\} = \exp\{(\lambda_1 + \lambda_2)(e^t - 1)\}$$

E então pode-se concluir que Y é Poisson com média $\lambda_1 + \lambda_2$

PROBLEMA 3 (25 pontos)

Suponha que X é uma variável aleatória contínua com densidade $f(x) = c/(1+x)^3$ onde $0 < x < 100$.

- Ache a constante c que faz de $f(x)$ uma densidade.
- Encontre a função de distribuição de X .
- Qual a probabilidade de X ser maior que 50?
- Encontre um ponto m no intervalo $(0,100)$ tal que $\Pr(X > m) = \Pr(X \leq m) = 50\%$. Este ponto é a *mediana* da distribuição.

Solução

a)

$$\int_0^{100} c \frac{1}{(1+x)^3} dx = 1 \Leftrightarrow c \frac{1}{(1+x)^2(-2)} \Big|_0^{100} = \frac{-c}{2} \left\{ \frac{1}{(101)^2} - 1 \right\} = 1 \Leftrightarrow \frac{c}{2} \left\{ 1 - \frac{1}{10201} \right\} = \frac{c}{2} \left\{ \frac{10200}{10201} \right\} = \frac{5100c}{10201} = 1$$

$$\Leftrightarrow c = \frac{10201}{5100} \approx 2.0002$$

b) $F(x) = 0$ se $x < 0$, $F(x) = 1$ se $x > 100$

$$F(x) = \int_0^x \frac{10201}{5100(1+t)^3} dt = \frac{-10201}{(2)5100(1+t)^2} \Big|_0^x = \frac{-10201}{10200} \left\{ \frac{1}{(1+x)^2} - 1 \right\} = \frac{10201}{10200} \left\{ 1 - \frac{1}{(1+x)^2} \right\} \text{ onde } 0 \leq x \leq 100$$

c) $\Pr(X > 50) = 1 - \Pr(X \leq 50) = 1 - F(50) =$

$$= 1 - \frac{10201}{10200} \left\{ 1 - \frac{1}{(1+50)^2} \right\} = 1 - \frac{10201}{10200} \left\{ 1 - \frac{1}{(51)^2} \right\} = 1 - \frac{10201}{10200} \left\{ \frac{2600}{2601} \right\} = 0.0003$$

d) $\Pr(0 < X < m) = F(m) = 50\%$

E aproveitando o item b):

$$\frac{10201}{10200} \left\{ 1 - \frac{1}{(1+m)^2} \right\} = \frac{1}{2} \Rightarrow \left\{ 1 - \frac{1}{(1+m)^2} \right\} = \frac{1}{2} \left(\frac{10200}{10201} \right) = \frac{5100}{10201}$$

$$1 - \frac{1}{(1+m)^2} = \frac{5100}{10201} \Rightarrow -\frac{1}{(1+m)^2} = \frac{5100 - 10201}{10201} = \frac{-5101}{10201} \Rightarrow \frac{1}{(1+m)^2} = \frac{5101}{10201}$$

$$\Rightarrow (1+m)^2 = \frac{10201}{5101} \Rightarrow 1+m = \sqrt{\frac{10201}{5101}} \approx 1.4141 \Rightarrow m \approx 0.4141$$

PROBLEMA 4 (25 pontos)

Você quer montar um portfólio com dois ativos. Calcule o retorno médio e o risco (desvio padrão do retorno do portfólio) EM FUNÇÃO DE α (a proporção do ativo A no portfólio) sob as seguintes condições:

Ativo A: retorno médio = 1%, d.p. retorno = 2%

Ativo B: retorno médio = 2%, d.p. retorno = 5%

- A correlação entre os dois ativos é nula; **(5 pontos)**
- A correlação entre os dois ativos é -0.2. **(5 pontos)**
- Ache α que produz o portfólio de mínimo risco na situação a) e na situação b) e calcule quais são os riscos, medidos pelos desvios padrões.
- Nas situações a) e b), como os desvios padrões do portfólio de mínimo risco se comparam ao de um portfólio com pesos iguais para os dois ativos? Qual redução no risco foi alcançada?

Solução

O **retorno médio** do portfólio é, **em qualquer condição**:

$$E(P) = \alpha \cdot E(A) + (1-\alpha) \cdot E(B) = 0.01(\alpha) + 0.02(1-\alpha)$$

A variância do retorno do portfólio é (em termos de α e ρ , o coeficiente de correlação):

$$\begin{aligned} VAR(P) &= \alpha^2 \cdot VAR(A) + (1-\alpha)^2 \cdot VAR(B) + 2\alpha(1-\alpha)\rho\sqrt{VAR(A)}\sqrt{VAR(B)} = \\ &= \frac{4}{(10)^4} \alpha^2 + \frac{25}{(10)^4} (1-\alpha)^2 + 2\rho\alpha(1-\alpha) \left(\frac{2}{100} \right) \left(\frac{5}{100} \right) = \\ &= \frac{1}{(10)^4} \{ 4\alpha^2 + 25(1-\alpha)^2 + 20\rho\alpha(1-\alpha) \} \end{aligned}$$

O risco do portfólio é o seu desvio padrão, isto é:

$$dp(P) = \sqrt{\frac{1}{(10)^4} \{ 4\alpha^2 + 25(1-\alpha)^2 + 20\rho\alpha(1-\alpha) \}}$$

a) Neste caso, $\rho = 0$ e o risco do portfólio é:

$$dp(P) = \sqrt{\frac{1}{(10)^4} \{4\alpha^2 + 25(1-\alpha)^2\}}$$

b) Neste caso, $\rho = -0.2$ e o risco do portfólio é:

$$dp(P) = \sqrt{\frac{1}{(10)^4} \{4\alpha^2 + 25(1-\alpha)^2 - 4\alpha(1-\alpha)\}}$$

c) Para minimizar o risco do portfólio, precisamos encontrar α que minimiza a variância (o que é equivalente a minimizar o desvio padrão do portfólio).

$$\frac{dVAR(P)}{d\alpha} = \frac{1}{(10)^4} \frac{d}{d\alpha} \{4\alpha^2 + 25(1-\alpha)^2 + 20\rho\alpha(1-\alpha)\} = 0$$

$$\Rightarrow 8\alpha + 25(-2 + 2\alpha) + 20\rho(1 - 2\alpha) = 0$$

$$\Rightarrow 8\alpha + 50\alpha - 40\rho\alpha = 50 - 20\rho$$

No caso a) $\rho = 0$, o valor de α que minimiza a variância do portfólio é então:

$$8\alpha + 50\alpha = 50 \Rightarrow 58\alpha = 50 \Rightarrow \alpha = 50/58 = 25/29$$

No caso b), $\rho = -0.2$, o valor de α que minimiza a variância do portfólio é então:

$$8\alpha + 50\alpha - 40\rho\alpha = 50 - 20\rho \Rightarrow 58\alpha + 8\alpha = 50 - 4 \Rightarrow 66\alpha = 46 \Rightarrow \alpha = 46/66 = 23/33$$

Na tabela abaixo os riscos de diversos portfólios, de acordo com diversos valores de α .

	$\rho = 0$	$\rho = -0.2$
$\alpha = 0$	5%	5%
$\alpha = 1/2$	2.69%	2.50%
$\alpha = 1$	2%	2%
Portfólio de mínimo risco	1.86%	1.84%

IND 1115 - Inferência Estatística

Profa. Mônica Barros

FORMULÁRIO P1

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	$f(x) = \frac{1}{2^{n/2} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}}$	n	2n	$\left(\frac{1}{1-2t}\right)^{n/2}$ se $t < 1/2$
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	$E(e^{tx}) = e^{\lambda \cdot (e^t - 1)}$
Binomial Negativa	$f(x) = \Pr(X = x) = \binom{x-1}{r-1} \cdot p^r \cdot q^{x-r}$ onde $x = r, r+1, r+2, \dots$	r/p	$r \cdot q/p^2$	$M(t) = \left(\frac{pe^t}{1-qe^t}\right)^r$

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

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

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

Relação entre densidade e função de distribuição

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

Definição: k-ésimo momento	Definição: Média ou Valor Esperado de X
$E(X^k) = \begin{cases} \int_{-\infty}^{\infty} x^k \cdot f(x) dx & \text{se } X \text{ é v.a. contínua} \\ \sum_{\text{todo } x} x^k \cdot f(x) = \sum_{\text{todo } x} x^k \cdot \Pr(X = x) & \text{se } X \text{ é v.a. discreta} \end{cases}$	$\mu = E(X) = \begin{cases} \int_{-\infty}^{\infty} x \cdot f(x) dx & \text{se } X \text{ é v.a. contínua} \\ \sum_{\text{todo } x} x \cdot f(x) = \sum_{\text{todo } x} x \cdot \Pr(X = x) & \text{se } X \text{ é v.a. discreta} \end{cases}$

Definição: k-ésimo momento central
$E((X - \mu)^k) = \begin{cases} \int_{-\infty}^{\infty} (x - \mu)^k \cdot f(x) dx & \text{se } X \text{ é v.a. contínua} \\ \sum_{\text{todo } x} (x - \mu)^k \cdot f(x) = \sum_{\text{todo } x} (x - \mu)^k \cdot \Pr(X = x) & \text{se } X \text{ é v.a. discreta} \end{cases}$
<i>Em particular, se $k = 1$: $E(X - m) = 0$, ou seja, o primeiro momento central é sempre nulo.</i>
Definição: Variância
$\sigma^2 = \text{VAR}(X) = E((X - \mu)^2) = \begin{cases} \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx & \text{se } X \text{ contínua} \\ \sum_{\text{todo } x} (x - \mu)^2 \cdot f(x) = \sum_{\text{todo } x} (x - \mu)^2 \cdot \Pr(X = x) & \text{se } X \text{ discreta} \end{cases}$

Fórmula alternativa para o cálculo da variância	Definição: Desvio padrão	Definição: Valor esperado de uma função de uma variável aleatória
$\sigma^2 = \text{VAR}(X) = E(X^2) - \mu^2$	$\sigma = \sqrt{\sigma^2} = \sqrt{\text{VAR}(X)}$	$E(u(X)) = \begin{cases} \int_{-\infty}^{\infty} u(x) \cdot f(x) dx & \text{se } X \text{ é v.a. contínua} \\ \sum_{\text{todo } x} u(x) \cdot f(x) = \sum_{\text{todo } x} u(x) \cdot \Pr(X = x) & \text{se } X \text{ é v.a. discreta} \end{cases}$

Propriedade – linearidade do valor esperado: $E\{a.u(X) + b.v(X)\} = a E\{u(X)\} + b E\{v(X)\}$

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 $\lambda = n.p$.

Função Geradora de Momentos (fgm)

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

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

- No caso contínuo

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

- 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 a variância de Y é:
$E(Y) = E\left(a_0 + \sum_{i=1}^n a_i X_i\right) = a_0 + \sum_{i=1}^n a_i \cdot E(X_i) =$ $= a_0 + \sum_{i=1}^n a_i \cdot \mu_i \quad \text{onde } \mu_i = E(X_i)$	$VAR(Y) = \sum_{i=1}^n a_i^2 \cdot VAR(X_i) = \sum_{i=1}^n a_i^2 \cdot \sigma_i^2$ <p>onde $VAR(X_i) = \sigma_i^2$</p>

E a **fgm de Y** é:

$$\begin{aligned} M_Y(t) &= E(e^{tY}) = E(e^{t(a_0 + a_1 X_1 + \dots + a_n X_n)}) \\ &= E(e^{ta_0} e^{ta_1 X_1} \dots e^{ta_n X_n}) = \\ &= e^{ta_0} E(e^{ta_1 X_1} \dots e^{ta_n X_n}) \end{aligned}$$

e como consequência da independência

$$\begin{aligned} &= e^{ta_0} E(e^{ta_1 X_1}) E(e^{ta_2 X_2}) \dots E(e^{ta_n X_n}) = \\ &= e^{ta_0} M_{X_1}(ta_1) M_{X_2}(ta_2) \dots M_{X_n}(ta_n) \end{aligned}$$

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

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

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

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) $\Gamma(1) = 0! = 1$

$$4) \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$