

IND 1115 – Inferência Estatística – Semestre 2004.02
Teste 1 – 05/10/2004
GABARITO

Problema 1 (15 pontos)

Uma gulosa professora de estatística é “fissurada” por trufas de chocolate. Em busca da trufa ideal, ela vai provando chocolates de maneira independente.

A probabilidade dela gostar de uma trufa que prova é 80%. Ela decide passear por um shopping, provando todas as trufas que encontra, e decide parar só ao encontrar a 5ª. trufa “maravilhosa” (para “desespero” da balança que tem em casa!).

Qual a probabilidade dela ter que:

- Provar 6 trufas até encontrar a 5ª. trufa maravilhosa?
- Ter que “sofrer”, provando 10 trufas, até encontrar a 5ª. trufa maravilhosa?
- Na média, quantas trufas ela vai ter que provar até encontrar a 5ª. trufa maravilhosa?

Solução

O modelo indicado é Binomial Negativo com $r = 5$ e $p = 0.8$ pois ela só pára quando encontra a 5ª. trufa maravilhosa.

Seja X o número de tentativas (trufas consumidas) até encontrar a 5ª. trufa maravilhosa.

$$a) \Pr(X = 6) = \binom{5}{4} (0.8)^5 (0.2)^1 = 5(0.2)(0.8)^5 = (0.8)^5 = 0.3277$$

$$b) \Pr(X = 10) = \binom{9}{4} (0.8)^5 (0.2)^5 = 0.0132$$

c) A média de uma variável Binomial Negativa é r/p , neste caso, $5/0.8 = 6.25$.

Problema 2 (15 pontos)

Seja X uma variável aleatória com função de probabilidade Poisson com média λ , isto é:

$$\Pr(X = x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!} \text{ onde } x = 0, 1, 2, \dots$$

Encontre a função geradora de momentos de X e, a partir dela, mostre que $E(X) = \lambda$.

Dica: expansão de Taylor da exponencial.

Solução

$$M(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} \cdot \exp(\lambda e^t) = \exp(-\lambda + \lambda e^t)$$

A primeira derivada da função geradora de momentos é:

$$\frac{dM(t)}{dt} = \frac{d}{dt} \left\{ e^{-\lambda} \cdot \exp(\lambda e^t) \right\} = e^{-\lambda} \lambda e^t \exp(\lambda e^t)$$

$$\left. \frac{dM(t)}{dt} \right|_{t=0} = e^{-\lambda} \lambda e^0 \exp(\lambda e^0) = e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \lambda = E(X)$$

Problema 3 (20 pontos)

Você trabalha numa empresa de consultoria. Apenas 10% dos projetos apresentados resultam num contrato. Calcule as seguintes probabilidades:

- De que o primeiro contrato acontecerá no 4^o. projeto apresentado.
- De que o 3^o. contrato fechado acontecerá no 6^o. projeto apresentado.
- Se você faz exatamente 15 apresentações de projeto num mês, qual a probabilidade de fechar 2 contratos?
- Quais são os modelos probabilísticos usados nos itens a), b) e c) – basta escrever os nomes e os valores dos parâmetros.

Solução

a) Neste caso X representa a tentativa (ou apresentação) em que ocorre o 1^o. fechamento de contrato e X é uma variável Geométrica com probabilidade $p = 0.1$. Logo:

$$\Pr(X = 4) = (0.9)^3 (0.1) = 0.0729$$

b) Aqui a variável de interesse é o número de apresentações até que o 3^a. contrato seja fechado, ou seja, trata-se de uma variável Binomial Negativa com parâmetros $r = 3$ e $p = 0.1$.

$$\Pr(X = 6) = \binom{5}{2} (0.9)^3 (0.1)^3 = 0.0073$$

c) Neste caso o número de chamadas é fixo a priori e portanto temos uma variável Binomial. X aqui representa o número de apresentações que resultaram em contratos fechados dentre as 15 apresentações realizadas num mês. Então X é Bin($n = 15$, $p = 0.1$).

$$\Pr(X = 2) = \binom{15}{2} (0.1)^2 (0.9)^{13} = 0.2669$$

d) Os modelos são, respectivamente, Geométrica ($p = 0.1$), NegBinomial ($r = 3$, $p = 0.1$) e Binomial ($n = 15$, $p = 0.1$).

Problema 4 (30 pontos)

Fez-se uma pesquisa de preços de roupas masculinas num shopping center. Uma amostra dos produtos existentes revela que o preço das calças é uma variável Normal com média R\$ 80 e desvio padrão R\$ 30. O preço das camisas é, por sua vez, uma variável Normal com média R\$ 60 e desvio padrão R\$ 25. A correlação entre os preços de calças e camisas é 0.6. Calcule as seguintes probabilidades:

- De um par de calças custar entre R\$ 60 e R\$ 95.
- De um par de calças custar entre R\$ 60 e R\$ 95 sabendo que uma camisa custa R\$ 75 nesta loja.
- De um par de calças custar entre R\$ 60 e R\$ 95 sabendo que uma camisa custa R\$ 50 nesta loja.
- Qual é a distribuição condicional dos preços das camisas sabendo que o preço das calças é R\$ 100?
- Qual é a distribuição condicional dos preços das camisas sabendo que o preço das calças é R\$ 70?
- Escreva a média condicional dos preços das camisas como função de x , o preço das calças na loja.

Solução

Seja X_1 o preço ds calças. Então X_1 é $N(80, (30)^2)$.

$$\begin{aligned} a) \Pr(60 < X_1 < 95) &= \Pr\left(\frac{60-80}{30} < \frac{X_1-80}{30} < \frac{95-80}{30}\right) = \\ &= \Pr\left(\frac{-2}{3} < Z < \frac{1}{2}\right) = \Phi\left(\frac{1}{2}\right) - \Phi\left(-\frac{2}{3}\right) = \Phi\left(\frac{1}{2}\right) - 1 + \Phi\left(+\frac{2}{3}\right) = \\ &= 0.6915 - 1 + 0.7475 = 0.4390 \end{aligned}$$

b) Sabe-se que, se X_1 e X_2 tem densidade Normal bivariada $N(\mu_1, \mu_2, \rho, \sigma_1^2, \sigma_2^2)$ então:

$$(X_1 | X_2 = x_2) \sim N\left(\mu_1 + \rho \cdot \left(\frac{\sigma_1}{\sigma_2}\right) \cdot (x_2 - \mu_2), \sigma_1^2 \cdot (1 - \rho^2)\right)$$

Neste caso:

$$\begin{aligned} (X_1 | X_2 = 75) &\sim N\left(80 + 0.6 \cdot \left(\frac{30}{25}\right) \cdot (75 - 60), (30)^2 \cdot (1 - 0.6^2)\right) = \\ &= N(80 + 10.8, (30)^2 \cdot (0.64)) = N(90.8, (30)^2 \cdot (0.64)) = N(90.8, (24)^2) \end{aligned}$$

$$\begin{aligned} \Pr(60 < X_1 < 95 | X_2 = 75) &= \Pr\left(\frac{60-90.8}{24} < Z < \frac{95-90.8}{24}\right) = \\ &= \Pr(-1.2833 < Z < 0.175) = \Phi(0.175) - \Phi(-1.2833) = \Phi(0.175) - 1 + \Phi(+1.2833) = \\ &= 0.5694 + 0.9003 - 1 = 0.4697 \end{aligned}$$

c) Neste caso:

$$\begin{aligned} (X_1 | X_2 = 50) &\sim N\left(80 + 0.6 \cdot \left(\frac{30}{25}\right) \cdot (50 - 60), (30)^2 \cdot (1 - 0.6^2)\right) = \\ &= N(80 - 7.2, (30)^2 \cdot (0.64)) = N(72.8, (24)^2) \end{aligned}$$

$$\begin{aligned} \Pr(60 < X_1 < 95 | X_2 = 50) &= \Pr\left(\frac{60-72.8}{24} < Z < \frac{95-72.8}{24}\right) = \\ &= \Pr(-0.5333 < Z < 0.925) = \Phi(0.925) - \Phi(-0.5333) = \Phi(0.925) - 1 + \Phi(+0.5333) = \\ &= 0.8225 + 0.7031 - 1 = 0.5256 \end{aligned}$$

d) A densidade condicional de X_2 dado X_1 é :

$$(X_2 | X_1 = x_1) \sim N\left(\mu_2 + \rho \cdot \left(\frac{\sigma_2}{\sigma_1}\right) \cdot (x_1 - \mu_1), \sigma_2^2 \cdot (1 - \rho^2)\right)$$

Se o preço das calças é R\$ 100 então:

$$(X_2 | X_1 = 100) \sim N\left(60 + 0.6 \cdot \left(\frac{25}{30}\right) \cdot (100 - 80), 25^2 \cdot (1 - 0.6^2)\right) = N(70, 25^2 (0.8)^2)$$

e) Se o preço das calças é R\$ 70 então:

$$(X_2 | X_1 = 70) \sim N\left(60 + 0.6 \cdot \left(\frac{25}{30}\right) \cdot (70 - 80), 25^2 \cdot (1 - 0.6^2)\right) = N(55, 25^2 (0.8)^2)$$

f) A média condicional do preço das camisas como função do preço das calças é:

$$E(X_2 | X_1 = x) = 60 + 0.6 \cdot \left(\frac{25}{30}\right) \cdot (x - 80) = 60 + \frac{1}{2}(x - 80) = \frac{x}{2} + 20$$

Problema 5 (20 pontos)

Um computador gera 8 números aleatórios uniformemente distribuídos no intervalo (0,1).

- Calcule a probabilidade de que o menor destes números será menor que 0.1.
- Calcule o valor esperado do menor destes números.
- Encontre a densidade do maior destes 8 números.
- Encontre o valor esperado do maior destes 8 números.
- 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.

Teorema

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

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 $y \rightarrow 0.1$, $t \rightarrow 0.9$ e se $y \rightarrow 0$, $t \rightarrow 1$. 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:

$$f(w) = \frac{\Gamma(9)}{\Gamma(1)\Gamma(8)} w^{8-1} (1-w)^{1-1} = \frac{8!}{0!7!} w^7 = 8w^7 \quad \text{onde } 0 < w < 1$$

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

Tabela – Função de Distribuição N(0,1)

z	$\Phi(z)$		z	$\Phi(z)$		z	$\Phi(z)$
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%		2,0100	97,78%			
0,9800	83,65%						
0,9900	83,89%						