

## PROBABILITAT, PROCESSOS ESTOCÀSTICS I ESTADÍSTICA

Solució de l'Examen final

22 de juny de 2012

1. En un sistema multiprocessador amb  $n$  processadors  $P_1, P_2, \dots, P_n$  i  $m$  memòries  $M_1, M_2, \dots, M_m$ , cada processador demana, independentment dels altres i amb igual probabilitat, accés a alguna memòria. Es demana:
- Sigui  $X$  la variable aleatòria que denota el nombre de processadors que demanen accés a la memòria  $M_1$ . Calcula la seva funció de probabilitat. De quina variable aleatòria es tracta? Quina és la probabilitat que la memòria  $M_1$  estigui ocupada (és a dir, accedida per al menys un processador)? Quin és el nombre mitjà de processadors que demanen accés a la memòria  $M_1$ ? Comenta el resultat.
  - Sigui  $Y$  la variable aleatòria que correspon al nombre de memòries ocupades, i  $I_j$ ,  $j = 1, 2, \dots, m$ , les variables aleatòries indicadores tals que  $I_j = 1$  si la memòria  $M_j$  està ocupada i  $I_j = 0$  altrament. Expressa  $Y$  en termes de les variables  $I_j$  i calcula el nombre mitjà de memòries ocupades (anomenat *ample de banda*  $B$  del sistema). A quins valors tendeix  $B$  quan  $m$  és molt gran (amb  $n$  finit) i quan  $n$  és molt gran (amb  $m$  finit)? Raona els resultats.
  - Si el nombre  $m$  de memòries és molt gran ( $m \rightarrow \infty$ ), i el temps d'accés a la memòria  $M_j$  és una variable aleatòria exponencial de paràmetre  $\lambda = j$ , calcula la probabilitat  $P(T > 1)$ , on  $T$  és un temps d'accés determinat (escollit al atzar). Comenta la lògica del resultat obtingut.
  - Per a  $m$  molt gran i donat que un determinat temps d'accés ha sigut  $T > 1$ , quina és la probabilitat que es tractés de la memòria  $M_1$ ?

**Solució:**

- (a) La probabilitat que el procesador  $P_i$ ,  $i = 1, 2, \dots, n$ , accedeixi a la memòria  $M_1$  és  $\frac{1}{m}$  (degut a l'equiprobabilitat). Per tant, el nombre de processadors que demanen accés a aquesta memòria és una variable aleatòria binomial amb paràmetres  $N = n$  i  $p = \frac{1}{m}$ :

$$P_X(x) = \binom{n}{x} \left(\frac{1}{m}\right)^x \left(1 - \frac{1}{m}\right)^{n-x}, \quad x = 0, 1, \dots, n.$$

Així, la probabilitat que la memòria  $M_1$  estigui ocupada és:  $P(X > 0) = 1 - \left(1 - \frac{1}{m}\right)^n$ , i el nombre mitjà de processadors que demanen accés a la memòria  $M_1$  és:  $E(X) = \frac{n}{m}$ , que es pot interpretar diguent que els  $n$  processadors "és distribueixen equitativament" entre les  $m$  memòries.

- (b) De les definicions resulta que  $Y = \sum_{j=1}^m I_j$ . Per tant,

$$B = E(Y) = E\left(\sum_{j=1}^m I_j\right) = \sum_{j=1}^m E(I_j) = mE(I_1) = mP(X > 0) = m\left[1 - \left(1 - \frac{1}{m}\right)^n\right],$$

que tendeix a  $m$  quan  $n \rightarrow \infty$ , i a  $n$  quan  $m \rightarrow \infty$ . En efecte, quan  $n$  és molt gran, la probabilitat que les  $m$  memòries estiguin ocupades és molt alta; mentre que, si  $m$  és molt gran, no hi haurà conflictes entre els processadors i cada un demanarà accés a una sola memòria.

(c) Per la fórmula de la probabilitat total,

$$\begin{aligned} P(T > 1) &= \sum_{j=1}^m P(T > 1 | M_j) P(M_j) = \frac{1}{m} \sum_{j=1}^m \int_1^{\infty} j e^{-jt} dt \\ &= \frac{1}{m} \sum_{j=1}^m e^{-j} \rightarrow \frac{1}{m} \left( \frac{1}{1 - e^{-1}} - 1 \right) = \frac{1}{m(e-1)} \rightarrow 0. \end{aligned}$$

Resultat obvi si pensem que el temps mitjà d'accés de la memòria  $M_j$  és  $\frac{1}{j} \rightarrow 0$ .

(d) Apliquem Bayes:

$$P(P_1 | T > 1) = \frac{P(T > 1 | M_1) P(M_1)}{P(T > 1)} = \frac{e^{-1} \frac{1}{m}}{\frac{1}{m(e-1)}} = 1 - e^{-1} = 0,6321.$$

2. En la superfície de cert material apareixen defectes en forma de rectangle d'amplada  $X$  i d'alçada  $Y$ , on  $(X, Y)$  és una variable bidimensional uniforme a la regió  $\Omega_{XY} = \{(x, y) : 0 \leq x, y \leq 6\}$ . Considerem també la variable aleatòria  $A$  corresponent a l'àrea dels rectangles.

- (a) Són  $X$  i  $Y$  independents? Troba els moments conjunts  $m_{r,s} = E(X^r Y^s)$ .  
 (b) Calcula el coeficient de correlació  $\rho$  entre  $A$  i  $X$ . Podem concloure'n alguna cosa sobre la independència entre  $A$  i  $X$ ?

*Indicació: pots fer servir els moments calculats a l'apartat (a).*

- (c) Dibuixa el domini de la variable  $(X, A)$ ,  $\Omega_{XA}$ , i calcula la seva funció de densitat conjunta,  $f_{XA}(x, a)$ .  
 (d) Calcula la millor estimació de  $X$  donada  $A$ ,  $\hat{X} = E(X|A = a)$ .

**Solució:**

- (a) Una justificació geomètrica: són independents ja que són variables uniformes en un domini quadrat.

Una justificació algebraica: Sabem que  $f_{XY}(x, y) = 1/36$ , llavors

$$f_X(x) = \int_0^6 f_{XY}(x, y) dy = \frac{1}{6}, \quad 0 < x < 6,$$

$$f_Y(y) = \int_0^6 f_{XY}(x, y) dx = \frac{1}{6}, \quad 0 < y < 6,$$

i  $f_X(x)f_Y(y) = f_{XY}(x, y)$ .

Els moments,

$$m_{r,s} = E(X^r Y^s) = E(X^r)E(Y^s) = \int_0^6 x^r f_X(x) dx \int_0^6 y^s f_Y(y) dy = \frac{6^{r+s}}{(r+1)(s+1)},$$

on la segona igualtat és certa ja que les variables  $X$  i  $Y$  són independents.

- (b) Podem escriure  $A = XY$ . Aleshores,

$$\begin{aligned} \rho &= \frac{Cov(X, A)}{\sigma(X)\sigma(A)} = \frac{E(XA) - E(A)E(X)}{\sqrt{E(X^2) - E(X)^2}\sqrt{E(A^2) - E(A)^2}} \\ &= \frac{m_{2,1} - m_{1,1}m_{1,0}}{\sqrt{m_{2,0} - m_{1,0}^2}\sqrt{m_{2,2} - m_{1,1}^2}} = \sqrt{\frac{3}{7}} \approx 0,65465. \end{aligned}$$

- (c) El domini de la variable  $X$  és  $\Omega_X = (0, 6)$ . Donada un  $x \in (0, 6)$ , tenim que  $0 \leq a \leq 6x$ . Per tant,

$$\Omega_{XA} = \{(x, a) : 0 \leq a \leq 6x \leq 36\}.$$

El dibuix apareix a la Figura 1.

El càlcul de  $f_{XA}(x, a)$  es pot fer de dues maneres:

- Tenim el següent canvi de variables,

$$\begin{cases} X &= X \\ A &= XY \end{cases}$$

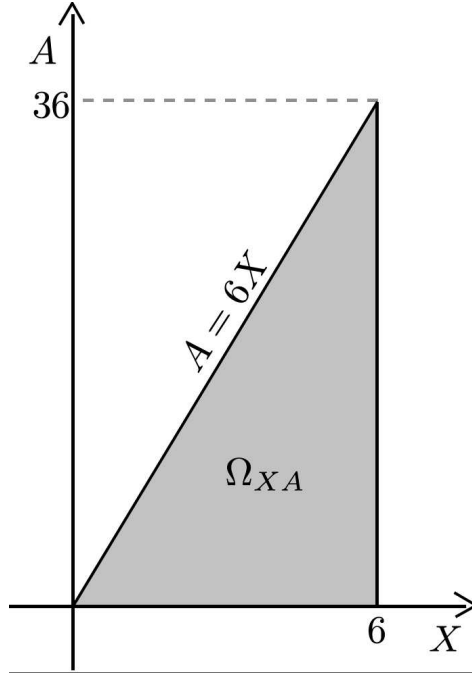


Figura 1: Domini  $\Omega_{XA}$ .

El determinant del Jacobià del canvi és,

$$\frac{\partial(x, a)}{\partial(x, y)} = \det \begin{pmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial a}{\partial x} & \frac{\partial a}{\partial y} \end{pmatrix} = \begin{vmatrix} 1 & 0 \\ y & x \end{vmatrix} = x.$$

Aleshores tenim que

$$f_{XA}(x, a) = f_{XY}(x, y) \frac{dx dy}{da} = f_{XY}(x, y) \left( \frac{\partial(x, a)}{\partial(x, y)} \right)^{-1} = \frac{1}{36x}.$$

- L'altre forma consisteix en calcular primer  $F_{XA}(x, a)$ ,

$$F_{XA}(x, a) = \Pr(X \leq x \cap A \leq a) =$$

$$\int_0^{a/6} \int_0^6 \frac{1}{36} dy dx + \int_{a/6}^x \int_0^{a/x} \frac{1}{36} dy dx = \dots = \frac{a}{36} (1 - \ln(a/6) + \ln x).$$

Per entendre millor els límits d'integració vegeu la Figura 2.

Per trobar la funció de densitat conjunta derivem  $F_{XA}(x, a)$  respecte  $x$  i  $a$ ,

$$f_{XA}(x, a) = \frac{\partial^2}{\partial a \partial x} F_{XA}(x, a) = \frac{1}{36x}.$$

- (d) Tenim que la millor estimació ve donada per

$$X \approx \tilde{X} = E(X | A = a) = \int_{a/6}^6 x f_X(x | A = a) dx = \int_{a/6}^6 x \frac{f_{XA}(x, a)}{f_A(a)} dx,$$

Podem calcular la marginal de  $A$ ,

$$f_A(a) = \int_{a/6}^6 f_{XA}(x, a) dx = \int_{a/6}^6 \frac{1}{36x} dx = \frac{1}{18} \ln 6 - \frac{1}{36} \ln a.$$

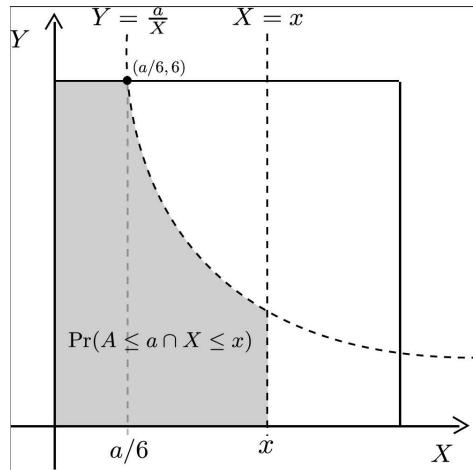


Figura 2: Càlcul de  $F_{XA}(x, a)$ .

Tenim que,

$$\tilde{X} = \int_{a/6}^6 x \frac{\frac{1}{36x}}{\frac{1}{18} \ln 6 - \frac{1}{36} \ln a} dx = \frac{1}{2 \ln 6 - \ln a} \int_{a/6}^6 dx = \frac{36 - a}{6(2 \ln 6 - \ln a)}.$$

La millor estimació de  $X$  en funció de  $A$  (Figura 3) és

$$X \approx \frac{36 - A}{6(2 \ln 6 - \ln A)}.$$

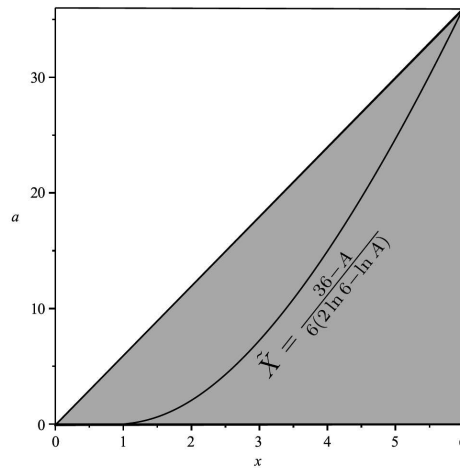


Figura 3: Millor estimació de  $X$  respecte  $A$ .

3. El nivell de càrrega en un acumulador d'energia ve donat pel procés estocàstic  $X(t) = At + B$ , per  $t \geq 0$ , on  $A$  i  $B$  són variables aleatòries uniformes en  $[0, 2]$ , independents.

- Calcula les funcions de valor mitjà, d'autocorrelació i d'autocovariància del procés  $X(t)$ . Es tracta d'un procés estacionari?
- Degut a fluctuacions ambientals un indicador de càrrega dona el valor  $Z(t) = X(t) + Y(t)$  on  $Y(t)$  és un procés de senyal telegràfic aleatori de paràmetre  $\lambda$ , independent del procés  $X(t)$ . Calcula les funcions de valor mitjà i d'autocorrelació de  $Z(t)$ . Com difereixen la potència mitjana real (la de  $X(t)$ ) i la potència mitjana en l'indicador?
- Calcula la millor estimació lineal no homogènia de  $X(t)$  donat  $Z(t)$ .
- Determina l'error quadràtic mínim en l'anterior estimació i fes-ne una gràfica comparant-lo amb l'error que es tindria si s'hagués estimat  $X(t)$  per una constant.

**Solució:**

$$(a) \quad E[A] = E[B] = \frac{0+2}{2} = 1, \quad E[A^2] = E[B^2] = V[A] + E[A]^2 = \frac{(2-0)^2}{12} + 1^2 = \frac{4}{3}, \\ E[AB] = E[A]E[B] = 1.$$

$$m_X(t) = E[X(t)] = E[A]t + E[B] = t + 1.$$

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = E[(At_1 + B)(At_2 + B)] \\ = E[A^2]t_1t_2 + E[B^2] + E[AB](t_1 + t_2) \\ = \frac{4}{3}t_1t_2 + \frac{4}{3} + t_1 + t_2.$$

$$C_X(t_1, t_2) = R_X(t_1, t_2) - m_X(t_1)m_X(t_2) = \frac{4}{3}t_1t_2 + \frac{4}{3} + t_1 + t_2 - (t_1+1)(t_2+1) = \frac{1}{3}(t_1t_2+1).$$

El procés no és estacionari ja que  $m_X(t)$  no és constant.

(b) Pel senyal telegràfic,  $m_Y(t) = 0$  i  $R_Y(t_1, t_2) = e^{-2\lambda|t_2-t_1|}$ .

$$m_Z(t) = E[Z(t)] = E[X(t) + Y(t)] = E[X(t)] + E[Y(t)] = t + 1.$$

$$R_Z(t_1, t_2) = E[Z(t_1)Z(t_2)] = E[(X(t_1) + Y(t_1))(X(t_2) + Y(t_2))] \\ = R_X(t_1, t_2) + R_Y(t_1, t_2) + m_X(t_1)m_Y(t_2) + m_Y(t_1)m_X(t_2) \\ = \frac{4}{3}t_1t_2 + \frac{4}{3} + t_1 + t_2 + e^{-2\lambda|t_2-t_1|}.$$

Pot $_X(t) = R_X(t, t) = \frac{4}{3}(t^2 + 1) + 2t$ . Pot $_Z(t) = R_Z(t, t) = \frac{4}{3}(t^2 + 1) + 2t + 1$ . La potència en l'indicador és una unitat superior a la potència real.

(c)  $\widehat{X}(t) = aZ(t) + b$ . Notem que  $E[X(t)Z(t)] = E[X(t)^2] + E[X(t)Y(t)] = \frac{4}{3}(t^2 + 1) + 2t$ . Pel principi d'ortogonalitat:

$$\begin{cases} E[Z(t)(aZ(t) + b)] = E[Z(t)X(t)] \\ E[1 \cdot (aZ(t) + b)] = E[1 \cdot X(t)] \end{cases} \\ \begin{cases} a(\frac{4}{3}(t^2 + 1) + 2t + 1) + b(t + 1) = \frac{4}{3}(t^2 + 1) + 2t \\ a(t + 1) + b = t + 1 \end{cases}$$

La solució és  $a = \frac{t^2 + 1}{t^2 + 4}$  i  $b = \frac{3(t + 1)}{t^2 + 4}$ .

(d) L'error quadràtic mitjà és:

$$\begin{aligned}\bar{\epsilon}_{\min} &= E[(X(t) - \widehat{X}(t))X(t)] = E[(X(t) - aZ(t) - b)X(t)] \\ &= \frac{4}{3}(t^2 + 1) + 2t - \frac{t^2 + 1}{t^2 + 4} \left( \frac{4}{3}(t^2 + 1) + 2t \right) - \frac{3(t + 1)}{t^2 + 4}(t + 1) = \frac{t^2 + 1}{t^2 + 4}.\end{aligned}$$

Estimant  $X(t)$  per una constant tenim un error igual la variància de  $X(t)$ , és a dir  $\bar{\epsilon}_{\min} = C_X(t, t) = \frac{1}{3}(t^2 + 1)$ . La gràfica següent mostra els dos errors.

