ETS d'Enginyeria de Telecomunicació de Barcelona

PROBABILITAT, PROCESSOS ESTOCÀSTICS I ESTADÍSTICA

Solució de l'Examen final **22 de juny de 2012**

---

1 Un canal transmet paraules d'un codi. Cada paraula consta de 4 bits que s'envien de forma independent. La probabilitat que un bit arribi canviat és $\frac{1}{3}$. Per tal de fer més segura la transmissió enviem cada bit dues vegades de forma independent. Direm que tenim un error en un bit si les dues transmissions d'aquest han fallat. Si N és la variable aleatòria que compta el nombre d'errors en un paraula, la probabilitat de recuperar la paraula és:
$P(\text{recuperar} \mid N = k) = 1 - \frac{k}{4}$ per a $0 \le k \le 4$.

(a) Doneu la probabilitat que una paraula transmesa es pugui recuperar.

(b) Quan el sistema ha rebut 4 paraules que no es poden recuperar, s'atura. Sigui T_i el nombre de paraules rebudes entre la $(i-1)$-èsima no recuperable i la i-èsima no recuperable (contant aquesta última). Justifiqueu que les variables T_i són independents i digueu de quin tipus de variable es tracten. El nombre de paraules transmeses quan el sistema s'atura és $T = T_1 + T_2 + T_3 + T_4$. Calculeu l'esperança i la variància de T.

(c) Quina és la probabilitat que les dues transmissions del primer bit diferexin entre si, si sabem que la paraula s'ha recuperat? Compareu-ho amb la probabilitat a priori i comenteu si la variació és raonable.

(d) Suposeu ara que rebem paraules de n bits i que la probabilitat d'error en un bit és 1/2. Acceptem una paraula si hem rebut com a molt $n/4$ errors. La refusem si te almenys $3n/4$ errors. Altrament, la tornem a demanar. Raoneu quina és la probabilitat que acceptem la paraula si sabem que no l'hem tornat a demanar. Quan n és molt gran, aproximeu la probabilitat de tornar-la a demanar en funció de n i de erf(). A què tendirà aquesta probabilitat quan $n \to +\infty$? Justifiqueu-ho.

Solució:

(a) Per cada bit de la paraula $P(\text{error}) = P(\text{fallen les dues transmissions}) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$. El nombre d'errors N és binomial amb $n = 4$, $p = \frac{1}{9}$.

$$\mathrm{P}(\text{recuperar}) = \sum_{k=0}^{4} P(\text{recuperar} \mid N = k) P(N = k)$$

$$= \binom{4}{0}\left(\frac{8}{9}\right)^4 + \binom{4}{1}\left(\frac{8}{9}\right)^3 \frac{1}{9} \cdot \frac{3}{4} + \binom{4}{2}\left(\frac{8}{9}\right)^2 \left(\frac{1}{9}\right)^2 \frac{1}{2} + \binom{4}{3}\frac{8}{9}\left(\frac{1}{9}\right)^3 \frac{1}{4} = \frac{8}{9}.$$

(b) Les variables T_i són geomètriques de paràmetre $p = 1/9$. Són independents ja que cada bit s'envia de forma independent. La variància de T_i val $\frac{q}{p^2} = 72$. Com són independents, la variància de la suma és suma de variàncies:

$$\mathrm{V}(T) = \mathrm{V}\left(\sum_{i=0}^{4} T_i\right) = \sum_{i=0}^{4} \mathrm{V}(T_i) = 4 \cdot 72 = 288.$$

(c) Utilitzant la fórmula de Bayes tenim,

$$\mathrm{P}(\text{1rs bits diferents} \mid \text{recuperada}) = \frac{\mathrm{P}(\text{recuperada} \mid \text{1rs bits diferents})\mathrm{P}(\text{1rs bits diferents})}{\mathrm{P}(\text{recuperada})}.$$

Aleshores, $\mathrm{P}(\text{recuperada}) = \frac{8}{9}$, $\mathrm{P}(\text{1rs bits diferents} \mid \text{recuperada}) = \frac{11}{12}\frac{4/9}{8/9} = \frac{11}{24}$,
$\mathrm{P}(\text{recuperada} \mid \text{1rs bits diferents}) = \binom{3}{0}\left(\frac{8}{9}\right)^3 + \binom{3}{1}\left(\frac{8}{9}\right)^2 \frac{1}{9} \cdot \frac{3}{4} + \binom{3}{2}\frac{8}{9}\left(\frac{1}{9}\right)^2 \frac{1}{2} + \binom{3}{3}\left(\frac{1}{9}\right)^3 \frac{1}{4} = \frac{11}{12}.$

$$\mathrm{P}(\text{1rs bits diferents} \mid \text{recuperada}) = \frac{11}{12}\frac{4/9}{8/9} = \frac{11}{24}.$$

Tenim que $11/24 > 4/9$. Si hem recuperat la paraula, vol dir que teniem pocs errors. Si sabem que no hi ha gaires errors, disminueix la probabilitat d'error en la primera posició. Si la primera posició no té errors, és més probable que els primer bits siguin diferents, que en el cas que no sapiguem res, ja que només tenim error en un bit si les dues transmissions fallen, que en particular implica que els primers bits són iguals.

(d)

$$\mathrm{P}(\text{acceptar}|\text{no la tornem a demanar}) = \mathrm{P}\left(N \leq \frac{n}{4} \,\middle|\, N \leq \frac{n}{4} \text{ o } N \geq \frac{3n}{4}\right)$$

$$= \frac{\mathrm{P}\left(N \leq \frac{n}{4}\right)}{\mathrm{P}\left(N \leq \frac{n}{4}\right) + \mathrm{P}\left(N \geq \frac{3n}{4}\right)} = \frac{1}{2}.$$

ja que amb $p = 1/2$ la funció de probabilitat és simètrica al voltant de $n/2$ i per tant $\mathrm{P}\left(N \leq \frac{n}{4}\right) = \mathrm{P}\left(N \geq \frac{3n}{4}\right)$.

(Alternativament, fixeu-vos que la probabilitat de rebre una paraula és la mateixa per totes les paraules binàries de longitud n, indiferentment de la paraula enviada. Podem aparellar una paraula P amb la seva complementària $\overline{P}$ que té tots els bits canviats. Observeu que si P te k errors, aleshores $\overline{P}$ en te $n-k$. Això implica que hi ha el mateix nombre de paraules que s'acceptaran que paraules que es refusaran. Com que totes tenen la mateixa probabilitat d'apareixer i sabem que no s'ha tornat a demanar la paraula, la probabilitat d'acceptar és $1/2$ indiferentment de la n.)

Sigui X el nombre de bits canviats. Aleshores $X \sim \text{Binomial}\left(n, \frac{1}{2}\right)$, $\mathrm{E}(X) = np = \frac{n}{2}$, $\sigma = \sqrt{np(1-p)} = \frac{\sqrt{n}}{2}$.

Aproximant X amb la gaussiana de distribució F_X:

$$\mathrm{P}(\text{retornar}) = \mathrm{P}\left(\frac{n}{4} < X < \frac{3n}{4}\right) = F_X\left(\frac{3n}{4}\right) - F_X\left(\frac{n}{4}\right)$$

$$= \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\frac{3n}{4} - \frac{n}{2}}{\sqrt{2}\frac{\sqrt{n}}{2}}\right)\right) - \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\frac{n}{4} - \frac{n}{2}}{\sqrt{2}\frac{\sqrt{n}}{2}}\right)\right) = \operatorname{erf}\left(\frac{\sqrt{n}}{2\sqrt{2}}\right).$$

Clarament, $\lim\limits_{n\to\infty} \mathrm{P}(\text{retornar}) = 1$. Sabem que una binomial està concentrada en una franja de longitud proporcial a $\sqrt{n}$ al voltant de la seva esperança, np. Per tant, la probabilitat de estar dins d'una franja de longitud $n/2$ al voltant de l'esperança tendeix a 1 ràpidament.

2 Un senyal electromagnètic es detecta en dos llocs diferents. En cada lloc hi ha un desfasament respecte el senyal original, donat per les variables aleatòries X i Y respectivament. La seva funció de densitat conjunta és:

$$f_{XY}(x,y) = \begin{cases} \dfrac{1}{4\pi^2}(1+\alpha \sin x \sin y), & \text{si } 0 \le x \le 2\pi,\ 0 \le y \le 2\pi \\ 0, & \text{altrament.} \end{cases}$$

on $\alpha \in [-1,1]$, és un paràmetre real.

(a) Trobeu les funcions de densitat marginal de les variables aleatòries X i Y. Quin tipus de variable són?

(b) Calculeu l'esperança condicionada $\mathrm{E}[Y|X=x]$. Dibuixeu la corresponent corba de regressió.

(c) Calculeu el coeficient de correlació ρ entre X i Y. Quin és el màxim valor que pot prendre? Són X i Y incorrelades per algun valor de α? En aquest cas: són independents?

(d) Calculeu el coeficient de correlació ρ entre $U = X+Y$ i $V = X-Y$. Són U i V incorrelades? Són U i V independents?

Solució:

(a)

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_0^{2\pi} \frac{1}{4\pi^2}(1+\alpha \sin x \sin y)dy$$
$$= \frac{1}{4\pi^2}\left(\int_0^{2\pi} dy + \alpha \sin x \int_0^{2\pi} \sin y dy\right) = \frac{1}{2\pi},$$

per a $0 \le x \le 2\pi$. Degut a la simetria $x \leftrightarrow y$, la variable Y té la mateixa densitat. X i Y són variables uniformes en l'interval $[0, 2\pi]$.

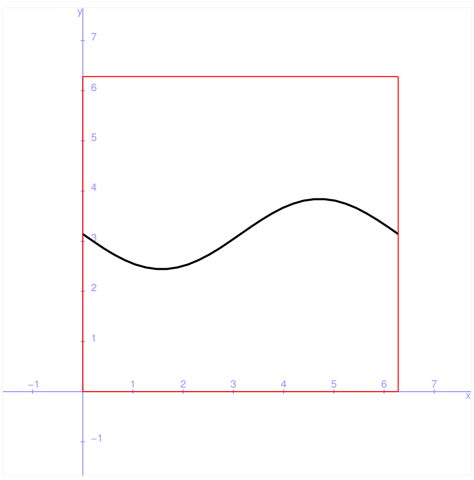
(b) La densitat de Y condicionada a X és (fixat $x \in [0, 2\pi]$)

$$f(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{1}{2\pi}(1+\alpha \sin x \sin y),$$

per a $0 \le y \le 2\pi$.

$$\mathrm{E}[Y|X=x] = \int_{-\infty}^{\infty} y f(y|x)dy = \frac{1}{2\pi}\int_0^{2\pi} y(1+\alpha \sin x \sin y)dy$$
$$= \frac{1}{2\pi}\left[\frac{y^2}{2} + \alpha \sin x(-y\cos y + \sin y)\right]_0^{2\pi} = \pi - \alpha \sin x.$$

La corba de regressió és $y = \pi - \alpha \sin x$ per a $0 \le x \le 2\pi$. Per $\alpha = 0{,}7$:



(c) Com són uniformes, $\mathrm{E}[X] = \mathrm{E}[Y] = \frac{0+2\pi}{2} = \pi$, $\sigma_X = \sigma_Y = \frac{2\pi-0}{\sqrt{12}} = \frac{\pi}{\sqrt{3}}$.

$$\begin{aligned}
\mathrm{E}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x,y)dxdy = \int_0^{2\pi}\int_0^{2\pi} xy\frac{1}{4\pi^2}(1+\alpha\sin x\sin y)dxdy \\
&= \frac{1}{4\pi^2}\left(\int_0^{2\pi} xdx\int_0^{2\pi} ydy + \alpha\int_0^{2\pi} x\sin xdx\int_0^{2\pi} y\sin ydy\right) \\
&= \frac{1}{4\pi^2}\left((2\pi^2)^2 + \alpha(-2\pi)^2\right) = \pi^2+\alpha,
\end{aligned}$$

així que $\mathrm{C}[X,Y] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] = \pi^2+\alpha-\pi^2 = \alpha$.

$$\rho = \frac{\mathrm{C}[X,Y]}{\sigma_X\sigma_Y} = \frac{3\alpha}{\pi^2}.$$

Els seus valors extrems són $\pm\frac{3}{\pi^2} = \pm 0{,}304$ (quan $\alpha = \pm 1$).

X i Y són incorrelades només si $\alpha = 0$. En aquest cas $f_{XY}(x,y) = \frac{1}{4\pi^2}$ i això coincideix amb $f_X(x)f_Y(y) = \frac{1}{2\pi}\cdot\frac{1}{2\pi}$, així que també són independents.

(d) $\mathrm{E}[U] = \mathrm{E}[X] + \mathrm{E}[Y] = 2\pi$, $\mathrm{E}[V] = \mathrm{E}[X] - \mathrm{E}[Y] = 0$, $\mathrm{E}[UV] = \mathrm{E}[(X+Y)(X-Y)] = \mathrm{E}[X^2-Y^2] = \mathrm{E}[X^2] - \mathrm{E}[Y^2] = 0$ ja que X i Y tenen la mateixa densitat. Llavors $\mathrm{C}[U,V] = \mathrm{E}[UV] - \mathrm{E}[U]\mathrm{E}[V] = 0$. Per tant, $\rho = 0$. Són incorrelades per a tot valor de α, però mai són independents ja que el quadrat on viuen X i Y es transforma girant 45 graus i la forma de rombe és incompatible amb la independència.

3 Sigui $X(t)$ el procés estocàstic

$$X(t) = \begin{cases} 2t, & \text{si } 0 \leq t \leq Y \\ 0, & \text{altrament,} \end{cases}$$

on Y és una variable aleatòria exponencial de paràmetre $\lambda = 1$.

(a) Quins són els possibles valors presos per cadascuna de les variables aleatòries $X(\frac{1}{2})$, $X(2)$ i $X(4)$. Determineu les seves funcions de probabilitat. Raoneu: es tracta d'un procés d'estat continu o d'estat discret?

(b) Calculeu la funció de probabilitat de primer ordre del procés $X(t)$. Utilitzeu-la per obtenir-ne la funció de valor mitjà $m(t)$, així com la potència del procés. Compareu la forma de les realitzacions amb la del valor mitjà. Quines similituds i diferències hi ha?

(c) Calculeu la funció d'autocorrelació de $X(t)$, $R(t_1, t_2)$. Verifiqueu a partir d'ella que la potència calculada a l'anterior apartat és correcta. *(Indicació: distingiu els casos $t_2 > t_1$ i $t_1 > t_2$.)*

(d) Determineu la millor estimació lineal no homogènia de $X(t_2)$ donat $X(t_1)$, $(0 < t_1 < t_2)$, i l'error quadràtic mínim d'aquesta estimació.

Solució:

(a) La variable aleatòria $X\left(\frac{1}{2}\right)$ pren el valors 0 i 1 amb probabilitats

$$\mathrm{P}\left(X\left(\frac{1}{2}\right) = 1\right) = \mathrm{P}\left(Y \geq \frac{1}{2}\right) = 1 - F_Y\left(\frac{1}{2}\right) = e^{-1/2},$$

$$\mathrm{P}\left(X\left(\frac{1}{2}\right) = 0\right) = 1 - \mathrm{P}\left(X\left(\frac{1}{2}\right) = 1\right) = 1 - e^{-1/2}.$$

Anàlogament, la variable aleatòria $X(2)$ pren els valors 0 i 4, mentre que $X(4)$ pren els valors 0 i 8. Les probabilitats respectives valen:

$$\mathrm{P}\left(X\left(2\right) = 4\right) = \mathrm{P}\left(Y \geq 2\right) = 1 - F_Y\left(2\right) = e^{-2},$$

$$\mathrm{P}\left(X\left(2\right) = 0\right) = 1 - \mathrm{P}\left(X\left(2\right) = 4\right) = 1 - e^{-2},$$

$$\mathrm{P}\left(X\left(4\right) = 8\right) = \mathrm{P}\left(Y \geq 4\right) = 1 - F_Y\left(4\right) = e^{-4},$$

$$\mathrm{P}\left(X\left(4\right) = 0\right) = 1 - \mathrm{P}\left(X\left(4\right) = 8\right) = 1 - e^{-4}.$$

En general, $X(t)$ és una variable aleatòria discreta que pren el valor $2t$ si $Y \geq t$ i el valor 0 altrament. $X(t)$ és, doncs, un procés d'estat discret.

(b) Noteu que per a $t < 0$ el procés $X(t)$ pren el valor 0 amb probabilitat 1. Per tant, si $t < 0$, $m_X(t)$ i la potència $P(t)$ valen trivialment 0.

Si $t > 0$, d'acord amb l'apartat anterior tenim:

$$\mathrm{P}(X(t) = 2t) = \mathrm{P}(Y \geq t) = 1 - F_Y(t) = e^{-t},$$

$$\mathrm{P}(X(t) = 0) = 1 - \mathrm{P}(X(t) = 2t) = 1 - e^{-t}.$$

Per tant, si $t > 0$,

$$m_X(t) = \mathrm{E}(X(t)) = 2t\,\mathrm{P}(X(t) = 2t) = 2t\,e^{-t}.$$

$$P(t) = \mathrm{E}(X^2(t)) = 4t^2\,\mathrm{P}(X(t) = 2t) = 4t^2\,e^{-t}.$$

(c) Pel que fa a la funció d'autocorrelació tenim (prenent $t_1, t_2 > 0$):

$$R_X(t_1, t_2) = \mathrm{E}(X(t_1)X(t_2)) = 4t_1t_2\,\mathrm{P}(Y \geq t_1, Y \geq t_2)$$

$$= 4t_1t_2\,\mathrm{P}\left(Y \geq \max\left(t_1, t_2\right)\right) = 4t_1t_2\, e^{-\max\left(t_1, t_2\right)}.$$

Per tant,

$$P(t) = R_X(t,t) = 4t^2\, e^{-t},$$

reobtenint així el resultat de l'apartat anterior.

La covariància val

$$C_X(t_1, t_2) = R_X(t_1, t_2) - m_X(t_1)m_X(t_2) = 4t_1t_2\ \left(e^{-\max\left(t_1, t_2\right)} - e^{-(t_1+t_2)}\right).$$

(d) Sigui

$$\widehat{X(t_2)} = \alpha X(t_1) + \beta.$$

Plantejant les equacions ortogonals:

$$0 = \mathrm{E}(X(t_2) - \alpha X(t_1) - \beta) = m_X(t_2) - \alpha m_X(t_1) - \beta,$$

$$0 = \mathrm{E}\left((X(t_2) - \alpha X(t_1) - \beta)X(t_1)\right) = R_X(t_1, t_2) - \alpha R_X(t_1, t_1) - \beta m_X(t_1),$$

obtenim

$$\alpha = \frac{C_X(t_1, t_2)}{C_X(t_1, t_1)} = \frac{4t_1t_2\ \left(e^{-t_2} - e^{-(t_1+t_2)}\right)}{4t_1^2\ \left(e^{-t_1} - e^{-2t_1}\right)} = \frac{t_2}{t_1}\ e^{-(t_2-t_1)}.$$

$$\beta = m_X(t_2) - \alpha m_X(t_1) = 2t_2\, e^{-t_2} - \frac{t_2}{t_1}\ e^{-(t_2-t_1)} 2t_1 e^{-t_1} = 0$$

Per tant,

$$\widehat{X(t_2)} = \frac{t_2}{t_1}\ e^{-(t_2-t_1)} X(t_1).$$

Pel que fa a l'error tenim

$$\begin{aligned}
\epsilon = \mathrm{E}\left(\left(X(t_2) - \widehat{X(t_2)}\right) X(t_2)\right) &= \mathrm{E}\left(\left(X(t_2) - \alpha X(t_1) - \beta\right) X(t_2)\right) \\
&= R_X(t_2, t_2) - \alpha R_X(t_1, t_2) - \beta m_X(t_2) \\
&= 4t_2^2\ e^{-t_2} - \frac{t_2}{t_1}\ e^{-(t_2-t_1)} 4t_1t_2\ e^{-t_2} = 4t_2^2\ e^{-t_2}\left(1 - e^{-(t_2-t_1)}\right).
\end{aligned}$$