MT5002 Sannolikhetsteori II tenta 2022-06-02

$\boxed{1}$ (a) Den sannolikhetsgenererande funktionen för en icke-negativ heltalsvärd stokastisk variabel X definieras som

$$g_X(t) := \sum_{k=0}^{\infty} t^k \, \mathbb{P}(X=k).$$ (2p)

(b) Låt X, Y vara icke-negativa och heltalsvärda, samt oberoende. Notera först att

$$g_{X+Y}(0) = \mathbb{P}(X+Y=0) = \mathbb{P}(X=0, Y=0)$$
$$\overset{\text{oberoende}}{=} \mathbb{P}(X=0)\,\mathbb{P}(Y=0) = g_X(0)\cdot g_Y(0)$$ (1p)

För $t \neq 0$ gäller $g_X(t) = \mathbb{E}[t^X]$, vilket ger

$$g_{X+Y}(t) = \mathbb{E}[t^{X+Y}] \overset{\text{ober.}}{=} \mathbb{E}[t^X]\,\mathbb{E}[t^Y] = g_X(t)\cdot g_Y(t).$$ (2p)

(c) Notera att $g_X(t)$ är en potensserie som konvergerar för $|t|<1$. Därmed är $g_X(t)$ deriverbar med väldefinierad derivata på $|t|<1$, och vi kan fortsätta att derivera. (2p)

Vi noterar att

$$g_X^{(n)}(t) = \sum_{k=n}^{\infty} k(k-1)(k-2)\dots(k-n+1)\, t^{k-n}\, \mathbb{P}(X=k)$$

För $t=0$ får vi

$$g_X^{(n)}(0) = n! \cdot \mathbb{P}(X=n). \qquad (*)$$ (2p)

Antag nu att $g_X(t) = g_Y(t)$ för $|t|<1$. Det följer från $(*)$ att

$$\mathbb{P}(X=n) = \frac{1}{n!} g_X^{(n)}(0) = \frac{1}{n!} g_Y^{(n)}(0) = \mathbb{P}(Y=n).$$ (1p)

Detta gäller för alla $n \geq 0$.

$\boxed{2}$ (a) Låt (X,Y) vara en diskret slumpvektor. För $x \in \mathbb{R}$ så att $\mathbb{P}(X=x)>0$ definierar vi den betingade fördelningen av Y givet $X=x$ som fördelningen av Y med avseende på sannolikhetsmåttet $\mathbb{P}(\cdot \mid X=x)$. Dvs fördelningen som ges av

$$\mathbb{P}(Y=y \mid X=x) = \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(X=x)} \quad \forall y \in \mathbb{R}. \qquad (2p)$$

(b) Låt (X,Y) vara en kontinuerlig slumpvektor med täthet $f_{X,Y}$ och $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$. För x så att $f_X(x)>0$ definierar vi den betingade fördelningen för Y givet $X=x$ som den fördelning som har täthet

$$f_{Y\mid X=x}(y) := \frac{f_{X,Y}(x,y)}{f_X(x)} \quad \forall y \in \mathbb{R}. \qquad (2p)$$

(c) Låt (X,Y) vara en diskret slumpvektor. Antag att $\mathbb{E}|Y|<\infty$. Per definition gäller

$$\mathbb{E}[Y \mid X=x] \overset{def}{=} \sum_{y \in \mathbb{R}} y \cdot \mathbb{P}(Y=y \mid X=x). \qquad (2p)$$

Sätt $h(x) := \mathbb{E}[Y \mid X=x]$. Då gäller

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}[h(X)] &\overset{def}{=} \sum_{x \in \mathbb{R}} h(x) \cdot \mathbb{P}(X=x) \qquad (2p)\\
&= \sum_{x \in \mathbb{R}} \sum_{y \in \mathbb{R}} y \cdot \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(X=x)} \cdot \mathbb{P}(X=x)\\
&\overset{!}{=} \sum_{y \in \mathbb{R}} y \sum_{x \in \mathbb{R}} \mathbb{P}(X=x, Y=y)\\
&= \sum_{y \in \mathbb{R}} y \cdot \mathbb{P}(Y=y) \qquad (2p)\\
&= \mathbb{E}[Y].
\end{aligned}$$

! Byte av summationsordning är tillåtet eftersom $\mathbb{E}|Y|<\infty$.

3 (a) Från tabell får vi att

$$\mathbb{E}[X] = \tfrac{1}{2} \quad \text{och} \quad \mathrm{Var}(X) = \tfrac{1}{12}.$$

Notera att $|Y| = X\,|2Z-1| = X$. Oberoende av X, Z ger

$$\mathbb{E}[Y] = \mathbb{E}[X]\left(2\,\mathbb{E}[Z]-1\right) = 0$$

$$\mathrm{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \mathbb{E}[X^2]$$

$$= \mathrm{Var}(X) + \mathbb{E}[X]^2 = \tfrac{1}{12} + \tfrac{1}{4} = \tfrac{4}{12} = 1/3.$$ (2p)

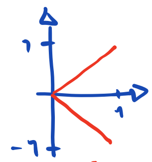
Dessutom har vi

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[XY]$$

$$= \mathbb{E}[X^2(2Z-1)] = \mathbb{E}[X^2]\left(2\,\mathbb{E}[Z]-1\right) = 0.$$ (2p)

Vi får

$$\Lambda = \begin{pmatrix} 1/12 & 0 \\ 0 & 1/3 \end{pmatrix}.$$ (1p)



(X,Y) tar värden på röd kurva

(b) Om X, Y oberoende så gäller speciellt att

$$\mathbb{E}[X \mid Y=y] = \mathbb{E}[X].$$ (1p)

Men $X = |Y|$ vilket för $y \neq \pm 1/2$ ger

$$\mathbb{E}[X \mid Y=y] = |y| \neq \mathbb{E}[X].$$

Alltså kan X och Y ej vara oberoende. (2p)

(c) I figur ovan så syns det att (X,Y) tar värden i en 1-dimensionell delmängd av planet, ty $Y = \pm X$. En dubbelintegral över ett område med area noll ger nödvändigtvis värde noll. Så ingen täthet kan existera. (2p)

4 (a) Notera att $Z = BX$ där

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Eftersom $X \sim N(0, I)$ så följer enligt sats att $Z \sim N(0, BIB')$, där

$$BB' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}.$$ (3p)

Vi noterar att

$$\det(BB') = 2\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} - 1\begin{vmatrix} 1 & 2 \\ -1 & 0 \end{vmatrix} = 2 - 2 = 0. \qquad (2p)$$

Täthet finns därför inte.

(b) Vektorn $(Z_1, Z_2) = (X_1, X_1 + X_2)$ är på nytt $N(0, \Delta)$-fördelad med $\Delta = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. $\det(\Delta) = 1$, så denna vektor har en täthet. Den betingade fördelningen av en komponent med avseende på en annan är igen normalfördelad. Vi behöver bestämma dess parametrar. (1p)

$$\mathbb{E}[Z_2 | Z_1 = x] = \mathbb{E}[X_1 + X_2 | X_1 = x] \overset{\text{ober.}}{=} x + \mathbb{E}[X_2] = x$$

$$\mathrm{Var}(Z_2 | Z_1 = x) = \mathbb{E}[(x + X_2 - \mathbb{E}[x + X_2 | X_1 = x])^2 | X_1 = x]$$

$$\overset{\text{ober.}}{=} \mathbb{E}[(x + X_2 - \mathbb{E}[x + X_2])^2]$$

$$= \mathrm{Var}(x + X_2)$$

$$= \mathrm{Var}(X_2) = 1. \qquad (2p)$$

Alltså är den betingade fördelningen $N(x, 1)$.

(c) Vi vill beräkna $\mathbb{E}[Z_2^2 Z_3^2]$. Från kovariansmatrisen ser vi att Z_2 och Z_3 är okorrelerade. Ty multi-normal är de då oberoende. Vi får

$$\mathbb{E}[Z_2^2 Z_3^2] \overset{\text{ober.}}{=} \mathbb{E}[Z_2^2]\,\mathbb{E}[Z_3^2] = \mathrm{Var}(Z_2)\mathrm{Var}(Z_3) = 4. \qquad (2p)$$

$\boxed{5}$ (a) Antag att $s_n^2 \xrightarrow{P} 1$ då $n \to \infty$. Enligt centrala gränsvärdessatsen gäller dessutom att

$$\sqrt{n}\,\bar{X} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \xrightarrow{d} N(0,1) \quad \text{då } n \to \infty. \qquad (2p)$$

Enligt Cramér–Slutsky gäller

$$T_n = \frac{\sqrt{n}\,\bar{X}}{\sqrt{s_n^2}} \xrightarrow{d} \frac{N(0,1)}{1} \quad \text{då } n \to \infty. \qquad (2p)$$

(b) Vi noterar att

$$s_n^2 = \frac{1}{n-1}\left[\sum_{k=1}^{n} X_k^2 - 2\sum_{k=1}^{n} X_k \bar{X} + n\bar{X}^2\right] = \frac{1}{n-1}\left[\sum_{k=1}^{n} X_k^2 - n\bar{X}^2\right]$$

$$= \frac{n}{n-1}\,\frac{1}{n}\sum_{k=1}^{n} X_k^2 - \frac{n}{n-1}\,\bar{X}^2 \qquad (2p)$$

Stora talens lag ger att

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mathbb{E}[X_k] = 0$$

$$\frac{1}{n}\sum_{k=1}^{n} X_k^2 \xrightarrow{P} \mathbb{E}[X_k^2] = 1$$ (1p)

Då $g(x)=x^2$ är kontinuerlig i $x=0$ följer det att $\bar{X}^2 \xrightarrow{P} 0$.

Cramér-Slutsky ger

$$\frac{n}{n-1}\cdot\left(\frac{1}{n}\sum_k X_k^2\right) \xrightarrow{d} 1\cdot 1 = 1$$

$$\frac{n}{n-1}\cdot\bar{X}^2 \xrightarrow{d} 1\cdot 0 = 0$$ (2p)

och därmed att

$$s_n^2 \xrightarrow{d} 1 \qquad \text{då } n\to\infty.$$

Eftersom konvergens i sannolikhet och fördelning är ekvivalent då gränsvärdet är konstant är vi klara. (1p)

$\boxed{6}$ (a) Från den sannolikhetsgenererande funktionen läser vi ut att Y har fördelning som ger sannolikhet $1/6$, $1/3$ och $1/2$ för utfallen $0, 1$ och 2. En sådan fördelning har väntevärde (2p)

$$\mathbb{E}[Y] = 0\cdot\frac{1}{6} + 1\cdot\frac{1}{3} + 2\cdot\frac{1}{2} = 4/3.$$ (1p)

(b) Enligt sats gäller av oberoendeantagandet att

$$\mathbb{E}[X_n] = \mathbb{E}\left[\sum_{k=1}^{X_{n-1}} Y_k^{(n)}\right] = \mathbb{E}[X_{n-1}]\cdot 4/3,$$ (2p)

givet att $\mathbb{E}[X_{n-1}]<\infty$. Upprepning ger

$$\mathbb{E}[X_n] = \left(\frac{4}{3}\right)^2\mathbb{E}[X_{n-2}] = \dots = \left(\frac{4}{3}\right)^n\cdot\mathbb{E}[X_0] = \left(\frac{4}{3}\right)^n.$$

Därmed gäller $\mathbb{E}[X_n]\to\infty$ då $n\to\infty$. (2p)

(c) Enligt sats gäller att $\eta := \mathbb{P}(T<\infty)$ är den minsta roten till ekvationen

$$t = g_Y(t).$$ (1p)

Denna ekvation blir

$$t = \frac{1}{6} + \frac{1}{3}t + \frac{1}{2}t^2$$

$$\Leftrightarrow \quad t^2 - \frac{4}{3}t + \frac{1}{3} = 0$$

$$\Leftrightarrow \quad \left(t - \frac{2}{3}\right)^2 + \frac{1}{3} - \frac{4}{9} = 0$$

$$\Leftrightarrow \quad t = \frac{2}{3} \pm \frac{1}{3}$$

Så $t = 1/3$ och $t = 1$ är lösningar. Alltså gäller

$$\mathbb{P}(T=\infty) = 1 - \eta = 1 - \frac{1}{3} = \frac{2}{3}.$$ (2p)