



Tentamen i matematisk statistik

Sannolikhetsteori A, 5 poäng

Skrivtid: 09.00-15.00.

Tillåtna hjälpmedel: Tabellsamling, egen miniräknare.
Studenterna får behålla tentamensuppgifterna.

- 1 På vissa cigarettpaket finns varningstexten ”9 av 10 som drabbats av strupcancer är rökare”. Denna faktauppgift är inte helt relevant för att bedöma om risken är stor eller liten för att drabbas av strupcancer, om man röker. Beräkna, med hjälp av informationen i varningstexten och följande data, sannolikheten att en rökare drabbas av strupcancer. 49% av befolkningen är män. Av männen röker 30% och av kvinnorna röker 40%. Sannolikheten att drabbas av strupcancer är 1%. (3 p)

Lösningförslag:

Låt C =händelsen cancer, R =händelsen rökare och M =händelsen man. Vi vet då att $P(C)=0.01$, $P(R|C)=0.9$, $P(M)=0.49$, $P(R|M)=0.3$, $P(R|M^*)=0.4$.

$$P(C|R) = \frac{P(C \cap R)}{P(R)} = \frac{P(R|C)P(C)}{P(R|M)P(M) + P(R|M^*)P(M^*)} = \frac{0.9 * 0.01}{0.3 * 0.49 + 0.4 * 0.51} = 0.025641$$

- 2 Låt (X, Y) vara en tvådimensionell s.v. med simultan frekvensfunktion

$$f(x, y) = \begin{cases} kx, & 0 < x < 1, x - 1 < y < 1 - x \\ 0, & \text{för övrigt} \end{cases}$$

- a) Bestäm värdet på konstanten k . (1 p)
b) Bestäm marginella frekvensfunktionen för Y . (1 p)
c) Bestäm variansen för X . (1 p)

Lösningförslag:

$$a) \quad 1 = \int_0^1 \int_{x-1}^{1-x} kx dy = \int_0^1 kx(1-x - (x-1))dx = \int_0^1 2kx(1-x)dx = k/3 \Rightarrow k = 3$$

b)

$$f_Y(y) = \begin{cases} \int_0^{y+1} 3x dx = \frac{3(y+1)^2}{2} & -1 < y \leq 0 \\ \int_0^{1-y} 3x dx = \frac{3(1-y)^2}{2} & 0 < y \leq 1 \end{cases}$$

$$c) \quad f_X(x) = \int_{x-1}^{1-x} 3x dy = 3x(1-x - (x-1)) = 6x(1-x), 0 < x \leq 1. \text{ Detta ger att } V(X)=1/20$$

- 3 Till en affär kommer kunder enligt en Poissonprocess med intensiteten 7 kunder per timme. Varje gång det kommer en kund så är det lika sannolikt att det är en man som en kvinna, oberoende av könet på tidigare kunder.
- a) Beräkna den betingade sannolikheten att det har kommit minst två män under den första halvtimmen, givet att det har kommit exakt fyra kunder sammanlagt under den första halvtimmen. (1.5 p)
- b) Beräkna väntevärdet av tiden det tar innan det kommer en kund av motsatt kön mot närmast föregående kund. (1.5 p)

Lösningsförslag:

Låt $X(t)$ =antal kunder under tiden $(0,t)$, $M(t)$ =antal män under tiden $(0,t)$ och $K(t)$ =antal kvinnor under tiden $(0,t)$. Då gäller att $\{X(t), t>0\}$ är en Poissonprocess med intensitet 7, $\{M(t), t>0\}$ och $\{K(t), t>0\}$ är oberoende Poissonprocesser med intensiteterna 3.5.

$$a) \quad P(M(0.5) \geq 2 | X(0.5) = 4) = 1 - P(M(0.5) = 0 | X(0.5) = 4) - P(M(0.5) = 1 | X(0.5) = 4) =$$

$$1 - \frac{P(M(0.5) = 0 \cap X(0.5) = 4)}{P(X(0.5) = 4)} - \frac{P(M(0.5) = 1 \cap X(0.5) = 4)}{P(X(0.5) = 4)} =$$

$$1 - \frac{P(M(0.5) = 0 \cap K(0.5) = 4)}{P(X(0.5) = 4)} - \frac{P(M(0.5) = 1 \cap K(0.5) = 3)}{P(X(0.5) = 4)} =$$

$$1 - \frac{P(M(0.5) = 0)P(K(0.5) = 4)}{P(X(0.5) = 4)} - \frac{P(M(0.5) = 1)P(K(0.5) = 3)}{P(X(0.5) = 4)} =$$

$$1 - \frac{e^{-1.75} * e^{-1.75} 1.75^4 / 4!}{e^{-3.5} 3.5^4 / 4!} - \frac{e^{-1.75} 1.75 * e^{-1.75} 1.75^3 / 3!}{e^{-3.5} 3.5^4 / 4!} =$$

$$1 - \left(\frac{1}{2}\right)^4 - 4\left(\frac{1}{2}\right)^4 = 1 - \frac{5}{16} = \frac{11}{16}$$

- b) Låt N vara numret på den första kund som har motsatt kön mot föregående kund. N kan då anta värdena $2, 3, 4, \dots$. Vi finner $P(N=2)=1*1/2$, $P(N=3)=1*1/2*1/2=(1/2)^2$ och att $P(N=k)=(1/2)^{k-1}$, $k=2, 3, \dots$. Av detta kan vi dra slutsatsen att $(N-1)$ är geometriskt fördelad med $p=1/2$ och att $E(N-1)=2$ dvs $E(N)=3$. Låt nu T_i vara tidpunkterna när kunder anländer, $i=1, 2, \dots$. Då gäller att T_i är exponentialfördelade med $E(T_i)=1/7$.

$$\text{Vi får då } E\left(\sum_{i=1}^N T_i\right) = E\left(E\left(\sum_{i=1}^N T_i\right) | N\right) = E(N) * E(T_i) = 3 * 1/7 = 3/7$$

- 4 Låt X_1 och X_2 vara två oberoende slumpstal med frekvensfunktion $f(x)=1, 0 \leq x \leq 1$.
- a) Bestäm frekvensfunktionen för X_1+X_2 och rita den i en figur. (0.75 p)
- b) Bestäm momentgenererande funktion för X_1+X_2 (0.75 p)
- c) Visa att X_1-X_2 har samma fördelning som X_1+X_2-1 (0.75 p)

- d) Bestäm frekvensfunktionen för $X_1 - X_2$ och rita den i en figur. (0.75 p)

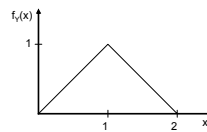
Lösningförslag:

- a) Låt $Y = X_1 + X_2$. Då

$$F_Y(y) = P(X_1 + X_2 \leq z) =$$

$$\iint_{\begin{cases} x \geq 0 \\ y \geq 0 \\ x+y \leq z \end{cases}} 1 dy dx = \begin{cases} \int_0^z \int_0^{z-x} 1 dy dx = \frac{z^2}{2} & 0 < z \leq 1 \\ \int_0^{z-1} \int_0^{z-1} 1 dx dy + \int_0^{z-1} \int_{z-1}^1 1 dx dy + \int_{z-1}^1 \int_{z-1}^{z-y} 1 dx dy = 1 - \frac{(2-z)^2}{2} & 1 < z \leq 2 \end{cases}$$

$$f_Y(x) = \begin{cases} z & 0 < z \leq 1 \\ 2-z & 1 < z \leq 2 \end{cases}$$



b) $M(Y, t) = M(X_1, t)M(X_2, t) = (E(e^{tX_1}))^2 = \left(\int_0^1 e^{tx} dx\right)^2 = \frac{1}{t^2}(e^t - 1)^2$

- c)

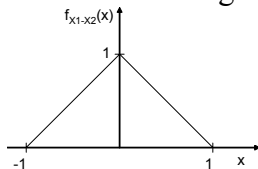
$$M(X_1 - X_2, t) = M(X_1, t)M(X_2, -t) = \frac{1}{t^2}(e^t + e^{-t} - 2)$$

$$M(X_1 + X_2 - 1, t) = e^{-t}M(X_1, t)M(X_2, t) = e^{-t} \frac{1}{t^2}(e^t - 1)^2 = \frac{1}{t^2}(e^t + e^{-t} - 2)$$

Eftersom mgf är unik måste $X_1 - X_2$ ha samma fördelning som $X_1 + X_2 - 1$

d) $F_{X_1 - X_2}(x) = P(X_1 - X_2 \leq x) = P(X_1 + X_2 - 1 \leq x) = P(X_1 + X_2 \leq x + 1)$

Genom derivering erhålles $f_{X_1 - X_2}(x) = f_Y(x + 1) = \begin{cases} x + 1 & -1 < x \leq 0 \\ 1 - x & 0 < x \leq 1 \end{cases}$



- 5 a) Definiera begreppet konvergens i sannolikhet för en följd av slumpvariabler $X_1, X_2, \dots$ (1 p)
b) Formulera Stora talens svaga lag (1 p)
c) Låt $P(A)$ beteckna sannolikheten att händelsen A inträffar vid ett försök. Gör n oberoende upprepningar av försöket och låt n_A vara antalet gånger A inträffar. Motivera att relativa frekvensen n_A/n konvergerar i sannolikhet mot $P(A)$. (1 p)
- 6 En belysningsarmatur består av tre lysrör som går sönder oberoende av varandra. Intensiteten (hazarden) som ett lysrör går sönder med är

$$h(t) = \begin{cases} 0.5, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

a) Bestäm den förväntade livslängden hos ett lysrör. (1.5p)

b) Anta att belysningen är tillräcklig om minst två lysrör lyser. Bestäm den förväntade tiden till dess att belysningen blir otillräcklig. (1.5p)

Lösningförslag:

a) X =livslängden har fördelningsfunktion $F(x) = 1 - e^{-\int_0^x 0.5 du} = 1 - e^{-x/2}$.

X är alltså exponentialfördelad med väntevärde 2.

b) S =systemets livslängd. $P(S > t) = P(2 \text{ av } 3 \text{ lysrör fungerar vid tid } t)$. Låt $p = P(X > t) = e^{-t/2}$ och Y =antal som fungerar vid tid t . Då $Y \in \text{Bin}(3, p)$ och $P(S > t) = P(Y=2) + P(Y=3) = 1 - P(Y=0) - P(Y=1) = 1 - (1-p)^3 - 3p(1-p)^2 = 1 - (1-e^{-t/2})^3 - 3e^{-t/2}(1-e^{-t/2})^2 = 3e^{-t} - 2e^{-3t/2}$. Vi erhåller

$$E(S) = \int_0^{\infty} (3e^{-t} - 2e^{-3t/2}) dt = 5/3$$

7 En Taylorutveckling av $g(X)$ kring $E(X)=m$ är

$$g(X) = g(m) + (X - m)g'(m) + \frac{(X - m)^2}{2!}g''(m) + \dots \text{ Vid Gauss approximation}$$

”kasseras” termer av ordning två och högre. Detta leder till uttrycken

$$E(g(X)) \approx g(m) \text{ och } V(g(X)) \approx V(X)g'(m)^2.$$

En bättre approximation uppnås om vi i Taylorutvecklingen också behåller 2:a grads-termen.

a) Härled på detta sätt ett approximativt uttryck för $E(g(X))$. (1 p)

b) Låt X vara en Poissonfördelad stokastisk variabel med väntevärde $m = 10$. Bestäm med vanlig Gauss-approximation approximationer av väntevärde och standardavvikelse för $g(X) = e^X$, samt med den nya varianten i a) en alternativ approximation av $E(e^X)$. (1 p)

c) Bestäm $E(e^X)$ exakt. Vad kan man säga om approximationerna i detta fall? (1 p)

Lösningförslag:

a) $E(g(X)) \approx g(m) + V(X)g''(m)/2$

b) Gauss: $E(e^X) \approx e^{E(X)} = e^{10} = 22026.46579$,
 $V(e^X) \approx V(X)e^{2m} = 10e^{20}$ och $D(e^X) \approx 69653.80070$
 Ny variant : $E(e^X) \approx e^{10} + 5e^{10} = 6e^{10} = 132158.7947$

c) $E(e^X) = \sum_{k=0}^{\infty} e^k \frac{e^{-10} 10^k}{k!} = \frac{e^{-10}}{e^{-10e}} \sum_{k=0}^{\infty} e^{-10e} \frac{(10e)^k}{k!} = e^{10(e-1)} = 2900034.487$

8 a) Visa Tjebysjevs olikhet, dvs

För varje $k > 0$ gäller $P(|X - \mu| \geq k\sigma) \leq 1/k^2$, där $E(X) = \mu$ och $V(X) = \sigma^2$.

Välj själv om du vill visa satsen för en diskret eller kontinuerlig s.v. (1.5 p)

b) Använd satsen för att utgående från n oberoende s v $X_1, X_2, \dots, X_n$ med samma fördelning som X , härleda ett konfidensintervall för μ med konfidensgrad minst 88,9% oavsett vilken fördelning X har. (3 p)

Lösningsförslag:

b) $E(\bar{X}) = \mu$ och $V(\bar{X}) = \frac{\sigma^2}{n}$. Enligt Tjebysjev's olikhet gäller då att

$P(|\bar{X} - \mu| > \frac{k\sigma}{\sqrt{n}}) \leq \frac{1}{k^2}$ eller att $P(|\bar{X} - \mu| \leq \frac{k\sigma}{\sqrt{n}}) > 1 - \frac{1}{k^2}$. Sätt $k=3$ ger

$P(|\bar{X} - \mu| \leq \frac{3\sigma}{\sqrt{n}}) > 0.889$ eller att $\bar{X} \pm 3 \frac{\sigma}{\sqrt{n}}$ är ett 88.9%-igt konfidensintervall för

μ .

Lycka till