STOCKHOLMS UNIVERSITET
Statistiska institutionen
Ellinor Fackle-Fornius

TENTAMEN I STATISTISK TEORI MED TILLÄMPNINGAR I

2017-10-26

Skrivtid: 15.00-20.00
Godkända hjälpmedel: Miniräknare, språklexikon.
Tentamen består av fem uppgifter. För full poäng på en uppgift krävs tydliga, utförliga och väl motiverade lösningar.
Resultatet meddelas senast den 9 november.

Uppgift 1. (20 poäng)

Enligt undersökningar av köpbeteendet hos kreditkortsinnehavare gäller följande simultana fördelning för X = antal ägda kreditkort och Y = antal kortköp per dygn.

		Y				
		0	1	2	3	4
	1	0.08	0.13	0.09	0.06	0.03
X	2	0.03	0.08	0.08	0.09	0.07
	3	0.01	0.03	0.06	0.08	0.08

a) Bestäm marginalfördelningen för antal kortköp per dygn.

b) Bestäm den betingade fördelningen för antal kortköp per dygn då antal ägda kreditkort är 3.

c) Vad är väntevärdet för antal kortköp per dygn för en person som äger 3 kreditkort?

d) Är variablerna antal ägda kreditkort och antal kortköp per dygn stokastiskt oberoende?

Uppgift 2. (20 poäng)

Tiden det tar för en kund att klippa sig hos frisörskan Hilda Hårfager kan beskrivas med en likformig fördelning i intervallet 30 till 80 minuter.

a) Bestäm fördelningsfunktionen för klipptiden i minuter.

b) Beräkna sannolikheten att en klippning tar mer än 60 minuter.

c) Hilda har under en dag 6 kunder. Antag att klipptiden för kunderna är oberoende. Vad är sannolikheten att minst 2 av dessa kunder behöver mer än 60 minuter för att bli färdigklippta?

d) Hilda tar betalt för en klippning baserat på en fast avgift samt en del som beror på hur lång tid klippningen tar. Konstnaden C för en kund är $C = 250 + 3.3Y$ där Y är klipptiden i minuter. Vad är väntevärdet och standardavvikelsen för kostnaden att klippa sig hos Hilda?

Uppgift 3. (20 poäng)

En viss typ av elektronisk komponent har en livslängd som är exponentialfördelad med väntevärde 2 år.

a) Beräkna sannolikheten att en sådan komponents livslängd överstiger 3 år.

b) Vad är median-livslängden för en sådan komponent?

c) En produkt innehåller 3 stycken av dessa elektroniska komponenter och fungerar till dess att någon av komponenterna går sönder. Bestäm fördelningsfunktionen för produktens livslängd då komponenterna kan antas vara oberoende.

d) Beräkna sannolikheten att livslängden för produkten i c)-uppgiften) överstiger 3 år.

Uppgift 4. (20 poäng)

Antag att den stokastiska variabeln U följer en Fréchet-fördelning med fördelningsfunktion

$$F(u) = \begin{cases} 0, & u \leq 0 \\ e^{-u^{-\alpha}}, & u > 0 \end{cases}$$

där parametern $\alpha > 0$. Fréchet-fördelningen används bland annat som modell inom extremvärdesteori, till exempel för att modellera maximal daglig nederbörd under en viss period.

a) Bestäm täthetsfunktionen för U.

b) Antag att Y är exponentialfördelad med $\beta = 1$. Visa att $U = Y^{-\frac{1}{\alpha}}$ är Fréchetfördelad.

Uppgift 5. (20 poäng)

Alfons och Bodil är båda två anställda i samma livsmedelsbutik. Låt Y_1 och Y_2 vara andelen tid (per arbetsdag) som Alfons repsektive Bodil sitter i någon av butikens kassor. Enligt en modell ges den simultana täthetsfunktionen för Y_1 och Y_2 av

$$f(y_1, y_2) = \begin{cases} y_1 + y_2, & 0 \leq y_1 \leq 1,\ 0 \leq y_2 \leq 1 \\ 0, & \text{annars.} \end{cases}$$

a) Bestäm marginalfördelningarna för Y_1 och Y_2.

b) Beräkna sannolikheten att både Alfons och Bodil sitter i kassan mer än halva arbetsdagen.

c) Beräkna sannolikheten att Alfons sitter i kassan mer än halva arbetsdagen givet att Bodil sitter i kassan mer än halva arbetsdagen.

Statistiska institutionen

Stockholms universitet

Rättningsblad

Datum: 26/10/17

Sal: Värtasalen

Tenta: Statistisk teori med tillämpningar I

Kurs: Statistisk teori med tillämpningar

ANONYMKOD: ST-0016

☐ Jag godkänner att min tenta får läggas ut anonymt på hemsidan som studentsvar.

OBS! SKRIV ÄVEN PÅ BAKSIDAN AV SKRIVBLADEN

Markera besvarade uppgifter med kryss

1	2	3	4	5	6	7	8	9	Antal inl. blad
X	X	X	X	X					6
Lär.ant. 5	13	10	16	16					

60 + 1 bonus

POÄNG	BETYG	Lärarens sign.
61	D	

SU, STATISTIK

Skrivsal: Värtasalen Anonymkod: ST-0016 Blad nr: 1

Uppgift 1

X = "antal ägda kreditkort"

Y = "antal kortköp per dygn"

a) Marginalfördelningen för antal kortköp per dygn erhålls genom att summera över Y.

X \ Y	0	1	2	3	4	
1	0,08	0,13	0,09	0,06	0,03	0,39
2	0,03	0,08	0,08	0,09	0,07	0,35
3	0,01	0,03	0,06	0,08	0,08	0,26
Σ	0,12	0,24	0,23	0,23	0,18	1

marginalfördelning för X.

Marginalfördelningen för antal kortköp per dygn (Y).

5

b) Betingad Slh-fördelning för Y givet X,

$$P(Y \mid X=3) = \frac{P(Y, X)}{P(X=3)} = \frac{1}{0{,}26}$$

0

c) Väntevärdet ges av:

$$E(X) = \sum_{x} x\, P(x)$$

Då $X = 3$: $E(X) = 3 \cdot 0{,}01 + 3 \cdot 0{,}03$

$+ 3 \cdot 0{,}06 + 3 \cdot 0{,}08 + 3 \cdot 0{,}08$

$= 0{,}78$

Svar: väntevärdet blir 0,78

d) Ja, variablerna X och Y är stokastiskt oberoende då, marginalfördelningarna för båda summerar till ett.

5

SU, STATISTIK

Skrivsal: Värta salen Anonymkod: ST-0016 Blad nr: 2

Uppgift 2

a) $Y =$ "tiden det tar att klippa sig i min"

$Y \sim Unif(30 - 80)$

Från formelblad får vi täthetsfunktionen:

$$f(y) = \begin{cases} \frac{1}{\theta_2 - \theta_1}, & \theta_1 \le y \le \theta_2 \\ 0, & \text{annars} \end{cases}$$

Då $\theta_2 = 80$ och $\theta_1 = 30$ har vi:

$$f(y) = \begin{cases} \frac{1}{50}, & 30 \le y \le 80 \\ 0, & \text{annars} \end{cases}$$

R

Fördelningsfunktionen får vi genom:

$$F(y) = \int_{30}^{y} \frac{1}{50}\,dt = \left[\frac{t}{50}\right]_{t=30}^{t=y} = \frac{y}{50} - \frac{30}{50}$$

R

Svar:

$$F(y) = \begin{cases} 0, & y < 30 \\ \frac{y}{50} - \frac{3}{5}, & 30 \le y \le 80 \\ 1, & y > 80 \end{cases}$$

R

4

b) Sannolikheten att en klippning tar mer än 60 min:

$$P(Y > 60) = 1 - P(Y \le 60)$$
$$= 1 - F(60) = 1 - \left(\frac{60}{50} - \frac{3}{5}\right) = 0,4$$

Svar: Sannolikheten att en klippning tar mer än 60 min är 0,4.

c) Vi har $P(Y > 60) = 0,4$

$$P = 0,4 \cdot \frac{2}{6} = \frac{2}{15} \approx 0,1333$$

d) För uniform fördelning: $E(Y) = \frac{\theta_1 + \theta_2}{2}$

$$E(Y) = \frac{30 + 80}{2} = 55$$

Väntevärde för kostnaden:

$$E(250 + 3,3Y) = E(250) + 3,3\,E(Y) =$$
$$= 250 + 3,3 \cdot 55 = 431,5$$

Standardavvikelsen för kostnaden:

$$V(Y) = \frac{(\theta_2 - \theta_1)^2}{12} = \frac{(80 - 30)^2}{12} = \frac{2500}{12}$$

$$V(250 + 3,3Y) = V(250) + 3,3^2 V(Y) = 250 + 3,3^2 \cdot \frac{2500}{12}$$
$$= 2518,75 \Rightarrow \sigma = \sqrt{2518,75} \approx 50,1871$$

SU, STATISTIK

Skrivsal: Värtasalen Anonymkod: ST-0016 Blad nr: 3

Uppgift 3

$Y =$ "livslängd i år"

a) $Y \sim \text{Exp}(\beta = 2)$

Från formelblad har vi att:

$$f(y) = \frac{1}{\beta} e^{-y/\beta}$$

Då $\beta = 2$ får vi: $f(y) = \frac{1}{2} e^{-y/2}$

Vi söker: $P(Y > 3) = 1 - P(Y < 3) = 1 - F(3)$

$$F(y) = \int_0^y \frac{1}{2} e^{-\frac{t}{2}} \, dt = \left[-e^{-\frac{t}{2}} \right]_{t=0}^{t=y} =$$

$$= -e^{-y/2} - (-e^0) = 1 - e^{-y/2}$$

Alltså blir: $P(Y > 3) = 1 - F(3) = 1 - \left(1 - e^{-3/2}\right)$

$$= e^{-3/2} \approx 0,2231$$

Svar: Sannolikheten att en sådan komponents livslängd överstiger 3 år är cirka 0,2231.

b) Vi söker median-livslängden för en sådan komponent.

Vi kan utnyttja att: $F(y) = 1 - e^{-y/2}$

$$1 - e^{-y/2} = 0,5$$

$$0,5 = e^{-y/2}$$

$$\ln(0,5) = -\frac{y}{2} \cdot \ln e$$

$$y = -2 \cdot \ln(0,5) \approx 1,3863$$

Svar: median-livslängden är cirka 1,3863 år.

c) Fördelningsfunktionen för en komponent:

$$F(y) = \begin{cases} 1 - e^{-y/2}, & y > 0 \\ 0, & \text{annars} \end{cases}$$

SU, STATISTIK

Skrivsal: Vårtasalen Anonymkod: ST-0016 Blad nr: 4

Uppgift 4

Vi har fördelningsfunktionen:

a)

$$F(u) = \begin{cases} 0, & u \leq 0 \\ e^{-u^{-\alpha}}, & u > 0 \end{cases}$$

Vi får täthetsfunktionen för U genom:

$$\frac{d}{du}\left[F(u)\right] = \frac{d}{du}\left[e^{-u^{-\alpha}}\right] = (-\alpha)(-u)^{-2\alpha} e^{-u^{-\alpha}} =$$

$$= \alpha u^{-2\alpha} e^{-u^{-\alpha}}$$

Svar:

$$f(u) = \begin{cases} \alpha u^{-2\alpha} \cdot e^{-u^{-\alpha}}, & u > 0 \\ 0, & \text{annars} \end{cases}$$

4

b) Vi har att $Y \sim Exp(\beta = 1)$ samt $U = Y^{-\frac{1}{\alpha}}$

Från formelblad får vi att:

$$f(y) = \frac{1}{\beta} e^{-y/\beta}, \quad \text{för } \beta > 0$$

$$0 < y < \infty$$

Då $\beta = 1$, har vi att:

$$f(y) = e^{-y}$$

Vi använder transformationsmetoden enligt:

$$f_U(u) = f_Y\left[h^{-1}(u)\right] \left|\frac{dh^{-1}}{du}\right|$$

$$h(u) = U = Y^{-\frac{1}{\alpha}} \Rightarrow Y = U^{-\alpha}$$

$$h^{-1}(u) = U^{-\alpha}$$

$$\frac{dh^{-1}}{du} = -\alpha U^{-2\alpha}$$

Alltså blir $f_U(u)$:

$$f_U(u) = e^{-u^{-\alpha}} \cdot \left|-\alpha U^{-2\alpha}\right| =$$

$$= \alpha U^{-2\alpha} e^{-u^{-\alpha}}$$

Svar: $U = Y^{-\frac{1}{\alpha}}$ är Frechetfördelad enligt:

$$f_U(u) = \begin{cases} \alpha U^{-2\alpha} e^{-u^{-\alpha}}, & u > 0 \\ 0, & \text{annars} \end{cases}$$

jmf. med täthetsfunktionen från uppgift a).

SU, STATISTIK

Skrivsal: Värtasalen Anonymkod: ST-0016 Blad nr: 5

Uppgift 5

Y_1 = "andelen tid Alfons sitter i kassan"

Y_2 = "andelen tid Bodil sitter i kassan"

$$f(y_1, y_2) = \begin{cases} y_1 + y_2, & 0 \le y_1 \le 1, \ 0 \le y_2 \le 1 \\ 0, & \text{annars} \end{cases}$$

a) Marginalfördelningarna för Y_1 och Y_2:

$$f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_2$$

$$f_1(y_1) = \int_0^1 (y_1 + y_2)\, dy_2 = \left[y_1 y_2 + \frac{y_2^2}{2} \right]_0^1 = y_1 + \frac{1}{2}$$

$$f_2(y_2) = \int_0^1 (y_1 + y_2)\, dy_1 = \left[\frac{y_1^2}{2} + y_2 y_1 \right]_0^1 = \frac{1}{2} + y_2$$

Svar:

$$f_1(y_1) = \begin{cases} y_1 + \frac{1}{2}, & 0 \le y_1 \le 1 \\ 0, & \text{annars} \end{cases} \qquad f_2(y_2) = \begin{cases} y_2 + \frac{1}{2}, & 0 \le y_2 \le 1 \\ 0, & \text{annars} \end{cases}$$

b) Vi ska beräkna Sth. att både Alfons och Bodil sitter i kassan mer än halva arbetsdagen.

$$P\left(Y_1 > \frac{1}{2}\right) = P\left(Y_2 > \frac{1}{2}\right)$$

Är dom beroende/oberoende?

Oberoende om $f(y_1, y_2) = f(y_1)\, f(y_2)$

Vi har att $y_1 + y_2 \neq \left(y_1 + \frac{1}{2}\right)\left(y_2 + \frac{1}{2}\right)$

Alltså är dom beroende!

Vi söker alltså:

$$\text{Sök: } \frac{P\left(Y_1 > \frac{1}{2} \cap Y_2 > \frac{1}{2}\right)}{P\left(Y_2 > \frac{1}{2}\right)} = \frac{\int_{1/2}^{1}\int_{1/2}^{1} f(y_1, y_2)\, dy_1\, dy_2}{\int_{1/2}^{1} f(y_2)\, dy_2}$$

$$\boxed{\begin{aligned} &\int_{1/2}^{1}\int_{1/2}^{1} y_1 + y_2 \, dy_1\, dy_2 = \int_{1/2}^{1} \left[\frac{y_1^2}{2} + y_2 y_1\right]_{1/2}^{1} dy_2 = \\ &= \int_{1/2}^{1} \left(\frac{1}{2} + y_2 - \frac{1}{4} - \frac{y_2}{2}\right) dy_2 = \int_{1/2}^{1} \frac{y_2}{2} + \frac{1}{4}\, dy_2 = \\ &= \left[\frac{y_2^2}{4} + \frac{y_2}{4}\right]_{1/2}^{1} = \frac{1}{4} + \frac{1}{4} - \frac{1}{16} - \frac{1}{8} = \frac{5}{16} \end{aligned}}$$