SDA III Tentamen 241009

1.

a) $\underline{X}$ = "antal hanterade skadeståndsansökningar till den första utbetalningen"

$\underline{X} \sim Geo(p=0.2)$, $p(x) = (1-p)^{x-1} p$ = / $p = 0.2$ / = $0.8^{x-1} \cdot 0.2$

$P(\underline{X}=4) = p(4) = 0.8^3 \cdot 0.2 = 0.1024$

Svar: Sannolikheten att en utbetalning sker efter att företaget har hanterat 4 ansökningar är ca 0.1.

4p.

b) $\underline{Y}$ = "antal skadeståndsansökningar som hanterats för att göra 5 utbetalningar"

$\underline{Y} \sim NegBin(r=5, p=0.2)$

$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$ = / $r=5$, $p=0.2$ / =

$\binom{y-1}{4} 0.2^5 \, 0.8^{y-5}$

$P(\underline{Y}=10) = p(10) = \binom{9}{4} 0.2^5 \, 0.8^5 = \frac{9!}{4! \, 5!} 0.2^5 \, 0.8^5$

$126 \cdot 0.2^5 \cdot 0.8^5 \approx 0.0132$

Svar: Sannolikheten att den 5:e utbetalningen sker efter att företaget har hanterat 10 ansökningar är ca 0.013.

5p.

c) När $Y \sim NegBin(5, 0.2)$ är väntevärdet och variansen

$$E(Y) = \frac{r}{p} = \frac{5}{0.2} = 25$$

$$V(Y) = \frac{r(1-p)}{p^2} = \frac{5 \cdot 0.8}{0.2^2} = 100$$

Svar: Väntevärdet och variansen för antal ansökningar som behöver hanteras för att genomför 5 utbetalningar är 25 respektive 100.

4p.

d) Administrativa kostnader:

700 kr / utbetalning

1600 kr / hanterad ansökan

Total administrativ kostnad för 5 utbetalningar:

$$C = 5 \cdot 700 + Y \cdot 1600 = 3500 + 1600\,Y$$

Sökt: $E(C)$ och $SD(C)$

$E(C) = E(3500 + 1600\,Y) = / E(c) = c,\ E(cY) = cE(Y) / =$

$3500 + 1600\,E(Y) = / E(Y) = 25 / = 3500 + 1600 \cdot 25 = 43\,500$

$V(C) = V(3500 + 1600\,Y) = / V(c) = 0,\ V(cY) = c^2 V(Y) / =$

$1600^2 V(Y) = / V(Y) = 100 / = 1600^2 \cdot 100$

$SD(C) = \sqrt{V(C)} = 1600 \cdot \sqrt{100} = 16000$

Svar: Väntevärdet och standardavvikelsen för att den totala administrativa kostnaden för att genomföra 5 utbetalningar är 43 500 kr respektive 16000 kr.

7p.

2.

a) X = "antal vinstlotter för Spel 1"

X har en diskret fördelning med följande sannolikhetsfunktion:

x	0	1	2	3
p(x)	0.4	0.3	0.2	0.1

Fördelningsfunktionen $F(x) = P(X \leq x)$ beräknas genom att summera de kumulativa slh:

$$F(0) = p(0) = 0.4$$

$$F(1) = p(0) + p(1) = 0.4 + 0.3 = 0.7$$

$$F(2) = p(0) + p(1) + p(2) = 0.4 + 0.3 + 0.2 = 0.9$$

$$F(3) = p(0) + p(1) + p(2) + p(3) = 0.4 + 0.3 + 0.2 + 0.1 = 1$$

vilket ger

$$\underline{\underline{F(x)}} = \begin{cases} 0, & x < 0 \\ 0.4, & 0 \leq x < 1 \\ 0.7, & 1 \leq x < 2 \\ 0.9, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

5p.

b) Y = "vinstsumma för Spel 2"

$Y \sim \text{Unif}(0, 100)$, $f(y) = \dfrac{1}{100}$, $0 < y < 100$

$$F(y) = \int_0^y \frac{1}{100}\,dt = \left[\frac{t}{100}\right]_{t=0}^{t=y} = \frac{y}{100} - 0 = \frac{y}{100}$$

dvs

$$F(y) = \begin{cases} 0, & y \leq 0 \\ \dfrac{y}{100}, & 0 < y < 100 \\ 1, & y \geq 1 \end{cases}$$

5p.

c) $U = Y^2$

Y har utfallsrum $0 < y < 100$, U är en strängt växande funktion i intervallet och har utfallsrum

$0 < u < 100^2 = 10\,000$

$$F_U(u) = P(U \leq u) = P(Y^2 \leq u) = P(Y \leq \sqrt{u}) =$$

$$F_Y(\sqrt{u}) = \frac{\sqrt{u}}{100}$$

$$f_U(u) = \frac{d}{du} F_U(u) = \frac{d}{du}\left(\frac{\sqrt{u}}{100}\right) = \frac{d}{du}\left(\frac{u^{1/2}}{100}\right) =$$

$$\frac{1}{2} \cdot \frac{u^{-1/2}}{100} = \frac{1}{200\sqrt{u}}$$

Sammanfattningsvis:

$$f_U(u) = \begin{cases} \frac{1}{200\sqrt{u}}, & 0 < u < 10\,000 \\ 0, & \text{annars} \end{cases}$$

$$F_U(u) = \begin{cases} 0, & u \leq 0 \\ \frac{\sqrt{u}}{100}, & 0 < u < 10\,000 \\ 1, & u \geq 10\,000 \end{cases}$$

10 p.

3.

a) Y_1 = " G.G.s aktieköp/mån i tiotusentals kr"

Y_2 = " G.G.s pokerspelvinst —ıı— "

$$f(y_1, y_2) = \begin{cases} k y_1 y_2, & 0 \leq y_1 \leq 1, \quad 0 \leq y_2 \leq 2 \\ 0, & \text{annars} \end{cases}$$

$$\int_0^2 \int_0^1 k y_1 y_2 \, dy_1 \, dy_2 = \int_0^2 \left[\frac{k y_1^2 y_2}{2} \right]_{y_1 = 0}^{y_2 = 1} dy_2 =$$

$$\int_0^2 \frac{k y_2}{2} \, dy_2 = \left[\frac{k y_2^2}{4} \right]_0^2 = k$$

Eftersom volymen under täthetsfunktionen är 1

har vi att $k = 1$

Svar: $k = 1$

6 p-

b) $$f_1(y_1) = \int_0^2 f(y_1, y_2) \, dy_2 = \int_0^2 y_1 y_2 \, dy_2 = \left[\frac{y_1 y_2^2}{2} \right]_{y_2 = 0}^{y_2 = 2} =$$

$2 y_1$, $0 \leq y_1 \leq 1$

$$f_2(y_2) = \int_0^1 f(y_1, y_2)\, dy_1 = \int_0^1 y_1 y_2\, dy_1 = \left[\frac{y_1^2 y_2}{2} \right]_{y_1=0}^{y_1=1} =$$

$$\frac{y_2}{2}, \quad 0 \leq y_2 \leq 2$$

Svar:

$$f_1(y_1) = \begin{cases} 2y_1, & 0 \leq y_1 \leq 1 \\ 0, & \text{annars} \end{cases}$$

$$f_2(y_2) = \begin{cases} \frac{y_2}{2}, & 0 \leq y_2 \leq 2 \\ 0, & \text{annars} \end{cases}$$

6p.

c) $$P(Y_1 > 0.5 \mid Y_2 < 1.2) = \frac{P(Y_1 > 0.5 \cap Y_2 < 1.2)}{P(Y_2 < 1.2)}$$

där

$$P(Y_1 > 0.5 \cap Y_2 < 1.2) = \int_0^{1.2} \int_{0.5}^{1} y_1 y_2\, dy_1\, dy_2 =$$

$$\int_0^{1.2} \left[\frac{y_1^2 y_2}{2} \right]_{y_1=0.5}^{y_1=1} dy_2 = \int_0^{1.2} \left(\frac{y_2}{2} - \frac{y_2}{8} \right) dy_2 = \int_0^{1.2} \frac{3y_2}{8}\, dy_2 =$$

$$\left[\frac{3y_2^2}{16} \right]_{y_2=0}^{y_2=1.2} = \frac{3 \cdot 1.2^2}{16} = 0.27$$

och

$$P(\underline{Y}_2 < 1.2) = \int_0^{1.2} \frac{y_2}{2}\, dy_2 = \left[\frac{y_2^2}{4} \right]_{y_2=0}^{y_2=1.2} = 0.36$$

vilket ger

$$P(\underline{Y}_1 > 0.5 \mid \underline{Y}_2 < 1.2) = \frac{0.27}{0.36} = 0.75$$

<u>Svar:</u> Slh att Gunnar köper aktier för mer än 5 tkr givet att hans pokervinst är mindre än 12 tkr är 0.75

8p.

4.

a) En prior är konjugerande för en modell när priorn och posteriorn tillhör samma fördelningsfamilj. Exempelvis är en normal prior konjugerande för normal likelihood då posteriorn också är normal, Beta är konjugerande för Bernoulli likelihood och Gamma för Poisson likelihood.

5p.

b) Ett Bayesian credible interval är ett intervall baserat på posterior fördelningen för parametern θ. Intervallet har tolkningen att slh att parametern θ ligger i det bayesianska intervallet är tex 95 %, för ett intervall som motsvarar 95 % av posterior-fördelningens slh-massa.

Inom klassisk inferens är konfidensintervallet konstruerat så att för ett 95 % igt k.i. kommer 95 % av ett stort antal k.i. beräknade på samma sätt att innehålla parametern θ. Slh att parametern θ ligger i ett uträknat k.i. är dock antingen 0 eller 1, då θ betraktas som fix. Istället används tolkas det som en viss grad av konfidens/säkerhet att intervallet innehåller θ.

5p.

5.

Y = "antal identifierade risker av 12 undersökta"

a) $Y \sim Bin(n=12, p=0.5)$ om H_0 är sann

$RR = \{y \leq 3\}$

$\alpha = P(\text{Forkasta } H_0 \text{ när } H_0 \text{ är sann}) =$

$P(Y \leq 3 \mid p = \frac{1}{2}) = p(0) + p(1) + p(2) + p(3) =$

$\text{pbinom}(3, 12, 0.5) \approx 0.073$

Svar: $\alpha = 0.073$

3p.

b) Testets styrka är

$P(\text{Förkasta } H_0 \text{ när } H_a \text{ är sann}) =$

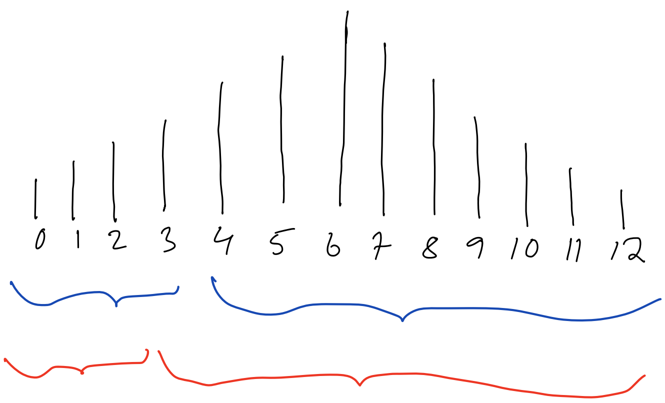
$P(Y \leq 3 \mid p = 0.3) = \text{pbinom}(3, 12, 0.3) \approx 0.49$

Svar: Styrkan är 0.49

3p.

c)

$p = 0.5$

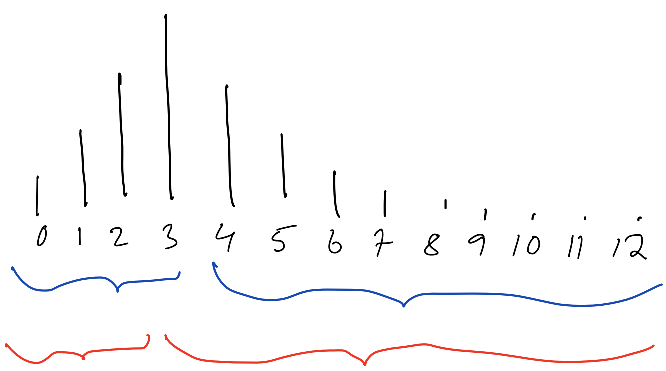


$RR = \{y \leq 3\}$ $AR = \{y \geq 4\}$

$RR = \{y \leq 2\}$ $AR = \{y \geq 3\}$

$\alpha \downarrow$ dvs risken att förkasta en sann nollhypotes minskar

$p = 0.3$



$RR = \{y \leq 3\}$ $AR = \{y \geq 4\}$

$RR = \{y \leq 2\}$ $AR = \{y \geq 3\}$

styrkan $\downarrow$ dvs slh att identifiera en falsk nollhypotes minskar också

4p.

6.

a) $f(y|\alpha) = \alpha y^{\alpha-1}$, $0 < y < 1$

Likelihoodfunktionen ges av, förutsatt slumpmässigt urval av n obs.

$$L(\alpha) = \prod_{i=1}^{n} f(y_i|\alpha) = \prod_{i=1}^{n} \alpha y_i^{\alpha-1} = \alpha^n \prod_{i=1}^{n} y_i^{\alpha-1}$$

och log likelihood

$$l(\alpha) = \ln[L(\alpha)] = \ln\left(\alpha^n \prod y_i^{\alpha-1}\right) =$$

$$\ln(\alpha^n) + \ln\left(\prod y_i^{\alpha-1}\right) = n\ln(\alpha) + \sum \ln(y_i^{\alpha-1})$$

$$n\ln(\alpha) + (\alpha-1)\sum \ln(y_i)$$

Derivera $l(\alpha)$, sätt lika med noll och lös ut α för att hitta $\hat{\alpha}_{ML}$:

$$\frac{\partial l(\alpha)}{\partial \alpha} = \frac{n}{\alpha} + \sum \ln(y_i)$$

$$\frac{n}{\alpha} + \sum \ln(y_i) = 0$$

$$\frac{n}{\alpha} = -\sum \ln(y_i)$$

$$\alpha = -\frac{n}{\sum \ln(y_i)}$$

<u>Svar:</u> $\hat{\alpha}_{ML} = -\frac{n}{\sum \ln(y_i)}$

b) För att testa

$H_0: \alpha = 2$

$H_a: \alpha \neq 2$

med ett LR-test används teststatistikan

$$\lambda_{LR} = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{\max_{\theta \in \Omega_0} L(\theta)}{\max_{\theta \in \Omega} L(\theta)}$$

som förkastas om λ_{LR} är tillräckligt liten

I detta fall har vi

$$\lambda_{LR} = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{L(\alpha = 2)}{L(\alpha = \hat{\alpha}_{ML})} = \frac{2^n \prod y_i}{\hat{\alpha}_{ML}^n \prod y_i^{\hat{\alpha}_{ML} - 1}}$$

När n är stort kan vi använda att

$$\bar{W} = -2\ln(\lambda_{LR}) \overset{\text{asymptotiskt}}{\sim} \chi^2(1)$$

$$\bar{W} = -2\ln\left[\frac{2^n \prod y_i}{\hat{\alpha}_{ML}^n \prod y_i^{\hat{\alpha}_{ML}-1}}\right] =$$

$$-2\left[n\ln(2) + \sum \ln(y_i) - n\ln(\hat{\alpha}_{ML}) - (\hat{\alpha}_{ML}-1)\sum\ln(y_i)\right] =$$

$$-2\left\{n\left[\ln(2) - \ln(\hat{\alpha}_{ML})\right] + \sum\ln(y_i)(2-\hat{\alpha}_{ML})\right\}$$

H_0 förkastas om $\bar{W} > \chi^2_{0.99}(1) = 6.635$

dvs $RR = \{\bar{W} > 6.635\}$