

LÖSNINGAR till tentamen: Statistik och sannolikhetslära (LMA120)

Tid och plats: Tisdagen den 3 juni, 8.30 - 12.30

Hjälpmedel: Typgodkänd miniräknare, formelsamling, kursbok samt anteckningar

Betygsgränser: 3: 12 poäng, 4: 18 poäng, 5: 24 poäng. Maximalt antal poäng: 30

Examinator: Petter Mostad

Telefonvakt: Petter Mostad

1. Beräkna $\binom{8}{3}$ och beskriv vad det är du har räknat ut. (2+2 poäng)

$\binom{8}{3}$ kan räknas ut antingen med hjälp av Pascals triangel eller definitionen. Om vi använder definitionen så får vi:

$$\binom{8}{3} = \frac{8!}{3!(8-3)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}{(1 \cdot 2 \cdot 3)(1 \cdot 2 \cdot 3 \cdot 4 \cdot 5)} = \frac{6 \cdot 7 \cdot 8}{1 \cdot 2 \cdot 3} = \frac{6 \cdot 7 \cdot 8}{6} = 7 \cdot 8 = 56$$

Vi har räknat ut på hur många olika sätt man kan välja 3 saker av totalt åtta saker, om ordningen inte spelar någon roll.

2. I en glassbar finns hundra olika glassmaker.

- (a) På hur många olika sätt kan vi välja två kulor glass, om ordningen inte spelar någon roll och vi kan välja samma smak två gånger?
(b) Om vi väljer de två kulorna helt slumpmässigt, vad är sannolikheten att de två kulorna vi väljer har samma smak? (3+3 poäng)

(a) Vi kan antingen välja två olika smaker, vilket vi kan göra på totalt $\binom{100}{2}$ sätt, eller välja samma smak två gånger, vilket vi kan göra på totalt $\binom{100}{1}$ sätt. Totalt kan vi alltså välja se två kulorna på

$$\binom{100}{2} + \binom{100}{1} = \frac{100 \cdot 99}{1 \cdot 2} + \frac{100}{1} = 5050$$

olika sätt.

(b) Observera först att det inte spelar någon roll vilken smak vi väljer först. När vi ska välja den andra smaken så finns det dock bara ett enda gynnsamt utfall, nämligen att välja samma smak som på den första kulan. Eftersom att det totalt finns 100 olika smaker så blir sannolikheten att vi väljer samma smak två gånger i rad därför $1/100 = 0.01$.

3. Antag att $X \sim N(0, 1)$. Beräkna följande sannolikheter:

- (a) $P(X \geq 2)$
(b) $P(X = 2)$
(c) $P(X > -2)$ (1+1+1 poäng)

(a) $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 2) = 1 - 0.9772 = 0.0228$ (se tabell)

(b) $P(X = 2) = 0$ eftersom att X är en kontinuerlig stokastisk variabel

(c) $P(X > -2) = P(X < 2) = 1 - P(X \geq 2) = 1 - 0.0228 = 0.9772$ enligt (a)

4. Antag att X är uniformt/likformigt fördelad på intervallet $[0, 1]$. Vad har $Y = 2X$ för fördelning? (3 poäng)

Eftersom att X är likformigt fördelad på intervallet $[0, 1]$ så har X kumulativ fördelningsfunktion

$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{om } x < 0 \\ x & \text{om } 0 \leq x \leq 1 \\ 1 & \text{om } x > 1 \end{cases}$$

För att ta reda på vad Y har för fördelning ska vi räkna ut den kumulativa fördelningsfunktionen till Y , dvs. räkna ut $P(Y \leq y)$:

$$P(Y \leq y) = P(2X \leq y) = P(X \leq y/2) = \begin{cases} 0 & \text{om } y < 0 \\ y/2 & \text{om } 0 \leq y/2 \leq 1 \\ 1 & \text{om } y/2 > 1 \end{cases} = \begin{cases} 0 & \text{om } y/2 < 0 \\ y/2 & \text{om } 0 \leq y \leq 2 \\ 1 & \text{om } y > 2 \end{cases}$$

Eftersom att det här är den kumulativa fördelningsfunktionen för en kontinuerlig stokastisk variabel som är likformigt fördelad på intervallet $[0, 2]$ så kan vi dra slutsatsen att Y måste ha denna fördelning.

5. Antag att $P(X = 1) = P(X = -1) = 0.5$. Beräkna $\mathbb{E}[X]$ och $\text{Var}(X)$.
(2+2 poäng)

$$\begin{aligned} \mathbb{E}[X] &= 1 \cdot P(X = 1) + (-1) \cdot P(X = -1) = 1 \cdot 0.5 - 1 \cdot 0.5 = 0 \\ \text{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - 0 \\ &= \mathbb{E}[X^2] \\ &= 1^2 \cdot P(X = 1) + (-1)^2 \cdot P(X = -1) \\ &= 1 \cdot 0.5 + 1 \cdot 0.5 \\ &= 1 \end{aligned}$$

6. Förklara vad som är fel i följande resonemang:

I en väljarundersökning frågade man 2000 personer om vilket parti de tänkte rösta på. Man beräknade sedan ett konfidensintervall med konfidensnivå 95% för andelen personer som tänker rösta på Piratpartiet i EU-valet och fick intervallet $[3.3 - 0.4, 3.3 + 0.4]$. Alltså är sannolikheten att fler än 3.7% av väljarna kommer att rösta på Piratpartiet minde än 2.5%.
(4 poäng)

För det första kan man aldrig från ett konfidensintervall för en parameter dra slutsatser om någon sannolikhet för parametern, till exempel sannolikheten för att parametern är utanför konfidensintervallet. I stället är det så att innan ett stickprov över huvud taget tagits, så är sannolikheten 95% att ett konfidensintervall med konfidensnivå 95% som man kan konstruera från dessa data innehåller parametern. För det andra stämmer längden av konfidensintervallet inte med konfidensnivån och antalet personer som har frågats: Om beräknat konfidensintervall är $[3.3, 3.7]$, så måste proportionen av de frågade som säger sig vilja rösta på Piratpartiet vara i mitten av intervallet, dvs. 3.5%, eller 0.035. Ett approximativt konfidensintervall med konfidensnivå 95% baserat på 2000 personer blir då

$$[0.035 - 1.96 \cdot \sqrt{0.035 \cdot (1 - 0.035)/2000}, 0.035 + 1.96 \cdot \sqrt{0.035 \cdot (1 - 0.035)/2000}] = [0.027, 0.043]$$

och alltså inte $[0.033, 0.037]$.

7. Vikten på ägg från en mataffär kan antas vara normalfördelad med okänt medelvärde μ och okänd standardavvikelse σ . För att ta reda på medelvikten hos äggen vägs alla ägg i en kartong med 12 ägg, och följande vikter erhålles (i gram):

52	47	61	37	37	51
55	50	53	49	48	59

- (a) Beräkna medianen, typvärdet och medelvärdet hos stickprovet. Vilket av de siffror du får tycker du säger mest om datan?
- (b) Beräkna stickprovsstandardavvikelsen s , givet att $\sum x_i^2 = 30493$, där x_i är vikten i gram hos ägg nummer i .

- (c) Beräkna ett konfidsensintervall för μ , med konfidensnivå $p = 0.98$.
(2+2+2 poäng)

- (a) Eftersom att det enda värde som förekommer två gånger är 37 så är typvärdet 37. Om vi sorterar vår data i storleksordning så får vi: 37 37 47 48 49 50 51 52 53 55 59 61. I mitten står de två värdena 50 och 51, så som median skulle vi kunna ta talet 50,5. Medelvärdet ges av

$$\bar{x} = \frac{37+37+47+48+49+50+51+52+53+55+59+61}{10} = 49,9$$

- (b) Vi räknar ut s^2 , stickprovsvariansen, först:

$$s^2 = \frac{\sum x_i^2 - n\bar{x}}{n-1} = \frac{30493 - 12 \cdot 49,9^2}{11} = 53,9$$

Vi får nu att

$$s = \sqrt{s^2} = \sqrt{53,9} = 7,34$$

- (c) Vårt konfidsensintervall ges av

$$\bar{x} \pm zs/\sqrt{n}$$

där vi redan har räknat ut s och $\bar{x}$. Det återstår därför att kontrollera i tabell vilket värde på z vi ska använda då $p = 0.98$. Vi har att $\alpha = 1 - p = 1 - 0.98 = 0.02$ och $df = n - 1 = 12 - 1 = 11$, så enligt tabellen för tvåsidiga konfidsensintervall ska vi använda $z = 2.72$. Vi får därför konfidsensintervallet

$$\bar{x} \pm zs/\sqrt{n} = 49,9 \pm 2.72 \cdot 7,34/\sqrt{11} = 49,9 \pm 6,02$$