

---

Lärare och Jour: David Bolin, telefon 772 53 75.

Tillåtna hjälpmedel: Formelsamling och valfri miniräknare med tömda minnen.

Korrekt, väl motiverad lösning ger poängen som är indikerad i parentes vid vardera uppgift. Totalt kan man få 40 poäng och betygsgränserna för betyg 3, 4 och 5 är 16, 24 och 32 poäng.

---

- (a)  $A$  och  $B$  är två händelser med  $P(A) = 0.3$ ,  $P(B|A) = 0.1$  och  $P(A \cup B) = 0.5$ . Bestäm  $P(B)$ . (2p)
  - (b)  $X$  och  $Y$  är två slumpvariabler med kovarians  $C(X, Y) = 0.5$  som uppfyller  $X \sim N(1, 2)$  och  $X + Y \sim N(3, 4)$ . Bestäm  $P(Y > 1)$ . (3p)
- En tvådimensionell slumpvariabel har täthetsfunktionen

$$f(x, y) = \begin{cases} \frac{(10-x)(5-y)}{C} & 0 \leq x \leq 10 \text{ och } 0 \leq y \leq 5, \\ 0 & \text{för övrigt.} \end{cases}$$

- Bestäm  $C$  så att  $f(x, y)$  blir en täthetsfunktion. (2p)
  - Beräkna den betingade fördelningen för  $X$  givet att  $Y = y$ . (2p)
  - Är  $X$  och  $Y$  oberoende? Motivera svaret. (1p)
- Folkhälsomyndigheten rekommenderar att temperaturen i lägenheter bör vara mellan 20 och 23 grader. Eftersom ett antal hyresgäster i en fastighet har klagat på låga temperaturer utför hyresvärden mätningar av temperaturen i 10 slumpvis utvalda rum, med följande resultat.

Rum $i$	1	2	3	4	5	6	7	8	9	10
Temp. ( $T_i$ )	19.35	21.18	19.24	18.89	19.15	19.43	19.44	20.18	19.80	20.59

- Antag att temperaturerna är normalfördelade och testa om medelvärdet understiger 20 grader. Kom ihåg att ange hypoteser och signifikansnivå. (3p)
  - Folkhälsomyndigheten skriver också att temperaturen aldrig ska vara lägre än 18 grader, och inte varmare än 26 grader på sommaren. Om temperaturen är normalfördelad med väntevärde 21.5 och standardavvikelse 2, vad är sannolikheten att temperaturen hamnar utanför det tillåtna intervallet  $[18, 26]$ ? (2p)
- En hyresgäst tvivlar på att analysen som hyresvärden gjorde i Uppgift 3 är korrekt, och hon vill istället att datan analyseras med hjälp av en icke-parametrisk metod.
    - Använd ett teckentest för att undersöka om mediantemperaturen understiger 20 grader, baserat på datan i Uppgift 3. Kom ihåg att ange hypoteser och signifikansnivå. (3p)
    - När bör ett teckentest användas istället för metoden som användes i Uppgift 3(a)? (1p)
    - En skillnad mellan testet i Uppgift 3 och teckentestet är att de har olika styrka. Vad är definitionen av ett hypotestests styrka? (1p)
  - Relativ ålderseffekt är ett välkänt fenomen inom många sporter. Detta uppkommer eftersom barn delas in i lag enligt födelseår, och att det därför kan skilja upp till 12 månader i ålder mellan barn inom samma grupp. Ett barn som är fött tidigt på året tenderar att vara mer moget och presterar därför bättre jämfört med de yngre barnen. Detta gör att dessa barn också tenderar att få mer uppmuntran och träning och därför fortsätter prestera bättre än sina yngre lagkamrater även i senare år. Resultatet av detta blir att topp-presterande idrottare

inom dessa sporter tenderar att vara födda tidigt på året. Kanada vann hockey-VM i år, och vi ska nu undersöka om det verkar finnas en relativ ålderseffekt inom hockey i Kanada. I tabellen nedan visas födelsedag för alla spelare i Kanadas nuvarande lag.

Namn	Födelsedag	Namn	Födelsedag
Taylor Hall	November 14	Cody Ceci	December 21
Ryan Ellis	Januari 3	Michael Matheson	Februari 27
Chris Tanev	December 20	Matt Duchene	Januari 16
Ben Hutton	April 20	Brendan Gallagher	Maj 6
Matt Dumba	Juli 25	Max Domi	Mars 2
Derick Brassard	September 20	Sam Reinhart	November 6
Corey Perry	Maj 16	Ryan Murray	September 27
Calvin Pickard	April 15	Cam Talbot	Juli 15
Boone Jenner	Juni 15	Morgan Rielly	Mars 9
Mark Scheifele	Mars 15	Mark Stone	Maj 13
Brad Marchand	Maj 11	Ryan O'Reilly	Februari 7
Connor McDavid	Januari 13		

- (a) Antag att det i snitt i befolkningen är lika sannolikhet att födas under de sex första månaderna på året som under de sex sista månaderna. Låt

$$X = \{\text{Antal spelare i laget som är födda under de sex första månaderna}\}.$$

Om fördelningen för födelsemånad för hockeyspelare är samma som för den övriga populationen, vilken fördelning borde då  $X$  ha? Motivera svaret. (1p)

- (b) I fördelningen anger parametern  $p$  sannolikheten att en hockeyspelare är född under de första sex månaderna av året. Härled uttrycket för maximum likelihood-skattningen  $p^*$  av parametern  $p$  och ange skattningen baserat på datan ovan. (4p)
- (c) Beräkna skattningens väntevärde  $E(p^*)$  och varians  $V(p^*)$ . (2p)
- (d) Utför ett hypotestest för att testa om  $p > 0.5$ . Använd signifikansnivå  $\alpha = 0.1$ , det är OK att använda normalapproximation. Vad är din slutsats, finns det bevis för en relativ ålderseffekt? (3p)

6. Sanna designar ett spökhäus inför Halloween och vill undersöka om mängden spindlar (Faktor A) samt mängden spindelnät (Faktor B) påverkar hur rädsla besökarna blir. Under tolv dagar utför hon ett  $2^2$ -faktorförsök med tre replikat, där responsvariabeln är medelantalet gånger besökarna skriker under besöket. Resultaten av undersökningen sammanfattas i följande tabell:

Faktor A	Faktor B	Respons		
Låg	Låg	1.35	3.18	1.24
Låg	Hög	4.89	5.15	5.43
Hög	Låg	3.44	4.18	3.80
Hög	Hög	8.59	7.15	8.80

- (a) Vad är en lämplig modell för experimentet? (1p)
- (b) Räkna ut de kvadratsummor som krävs för att beräkna variansanalystabellen för experimentet. (4p)
- (c) Använd resultatet från (b) för att beräkna variansanalystabellen. (2p)
- (d) Testa om det finns en signifikant samspelseffekt mellan faktorerna, på nivå  $\alpha = 0.05$ . (1p)
- (e) Testa om huvudeffekterna är signifikanta på nivå  $\alpha = 0.05$  (2p)

**Lycka till!**