
Korrekt och väl motiverad lösning på uppgifterna 1–5 ger 10 poäng vardera medan delfrågorna på uppgift 6 ger 4 poäng vardera. Totalt kan man få 70 poäng. Gränsen för godkänd är 35 poäng, dock finns det vissa minimikrav på uppgifterna 1–5 (18p) respektive uppgift 6 (7p).

Institutionens papper används både som kladdpapper och inskrivningspapper. Varje lösning skall börja överst på nytt papper. Rödpenna får ej användas.

Tillåtna hjälpmedel: Miniräknare, Formelsamling i matematisk statistik för B, K, N, BME och kemister och tabellsamling

Resultatet förs in i LADOK senast måndag den 19 november.

Redovisa införda beteckningar; ange modeller, approximationer, hypoteser och slutsatser. Motivera alla antaganden.

Skriv anonymkod och identifierare (eller personnummer och namn) på omslaget och SAMTLIGA inlämnade papper.

1. Två material till skosulor, A och B, testas på pojkar som får ha en sko med material A i sulan på en slumpmässigt vald fot och en sula med material B på den andra. Efter ett antal dagar mäts slitaget på de båda skorna, och resultatet visas i tabellen nedan.

Pojke nummer	A	B	B-A
1	13.2	14.0	0.8
2	8.2	8.8	0.6
3	10.9	11.2	0.3
4	14.3	14.2	-0.1
5	10.7	11.8	1.1
6	6.6	6.4	-0.2
7	9.5	9.8	0.3
8	10.8	11.3	0.5
9	8.8	9.3	0.5
10	13.3	13.6	0.3

Gör lämpliga antaganden om fördelning och oberoende.

- (a) Gör ett 95 procents konfidensintervall för skillnaden mellan materialen. (5p)
- (b) Drag ur svaret på första frågan en slutsats om skillnaden i slitage mellan materialen. (3p)
- (c) Någon föreslår att man istället ska se mätningarna från skorna med A-sulor respektive B-sulor som två stickprov och analysera enligt metoden "två stickprov". Förklara varför detta är en dålig idé! (2p)

(Ursprungligen ett exempel från Statistics For Experimenters , av Box, Hunter och Hunter, Wiley 1978.)

2. Kritikern François-Marie Arouet skriver litteraturrecensioner på tidningen Fäderneslandet. Han är en uppmärksam och eftertänksam kritiker, som dock inte helt kan skilja sina privata åsikter från mer objektiva kvalitetskriterier.

Han är mottaglig för smicker, och stadens unga författare har upptäckt att det kan löna sig ordentligt att bjuda med François-Marie på en krogrunda strax före ett boksläpp. Det har gått så långt att han ger en strålande recension till 25 procent av alla de usla diktsamlingar han får i uppgift att recensera. Bland de medelmåttiga diktsamlingarna ger han hela 30 procent strålande hyllningar med ord som "epokgörande" och "skolbildande för sin generation". De samlingar som verkligen

skulle förtjäna så ampla lovord får det också, men bara med sannolikhet 80 procent, eftersom det ibland händer att François-Marie tycker boken ligger för nära den bok han själv skulle ha velat skriva, och då hindrar avundsjukan honom från att berömma den.

Bland alla de diktsamlingar han recenserar är 30 procent usla, 68 procent medelmåttiga och 2 procent strålande och epokgörande.

- (a) Välj en slumpmässig diktsamling som recenserats av François-Marie. Vad är sannolikheten att den får strålande recensioner? (5p)
- (b) Du läser en strålande recension av diktsamlingen *Jag var en kärlekskonstens mästare* av J. B. Delgada utan att veta något om boken, men du ser att recensenten heter François-Marie Arout. Vad är sannolikheten att boken *verkligen* är strålande? (5p)
3. I Expressen 15/8 kan man läsa om projektet #slutaskjut för att minska gängbrottsligheten i Malmö. Tanken är att brottsligheten ska minska när de kriminella förstår konsekvenserna av sitt handlande.

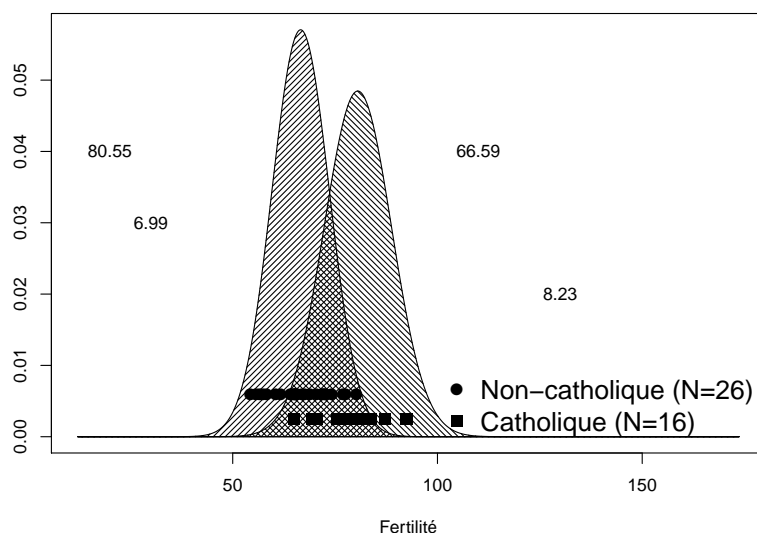
Tio av Malmös gängmän kallades till möte med polis och samhällets representanter. Nio av dem kom – och fick höra skakande vittnesmål om våldets konsekvenser.

– En mamma berättade om sorgen som föräldrar till mördade barn ständigt känner, inte ett öga var torrt, säger polischefen Stefan Sintéus.

Sei Shōnagon är polis i Malmö och tveksam, men vill ge projektet en chans. Därför rekryterar hon 400 slumpvis utvalda Malmöbor. 200 får rapportera om de varit i närheten av eller utsatts för något våldsbrott under ett utvalt veckoslut före projektets genomförande. 200 får rapportera samma sak ett veckoslut efteråt. Resultatet blev att 20 av dem som tillfrågats före projektet och 25 av dem efter svarade ”Ja”, vilket knappast gör Sei mindre tveksam.

Gör ett 95 procent tvåsidigt approximativt konfidensintervall för skillnaden i sannolikheten att utsättas/vara i närheten av våldsbrott före respektive efter #slutaskjut. Redovisa lämpliga antaganden om fördelning och oberoende! Vad kan Sei dra för slutsatser ur sitt försök? (10p)

(Det ska tydligt påpekas att data är påhittade. Hur projektet verkligen lyckas vet ingen än.)



4. Margaretha Geertruida Zelle rotar igenom några gamla papper hos sin döda schweiziska släkting Gottfried Keller och upptäcker att denne mot slutet av sitt liv ägnade sig åt fertilitetens koppling till religion. Hon förstår av texten att talen i figuren ovan beskriver ett mått på antal barn per kvinna i katolska respektive icke-katolska byar i det fransktalande Schweiz. Låt oss för enkelhets skull kalla talen från de katolska byarna y_i och dem från de ickekatoiska x_i .

Genom att läsa i texten förstår Geertruida att man utöver anpassade normalfördelningskurvor och antal byar i varje grupp kan se medelvärdena \bar{y} , \bar{x} och stickprovsstandardavvikelserna s_y , s_x i bilden. Vilket som är vad har dock fallit bort.

Data har skalats om så att den enhet fertiliteten mäts i är oklar. Men det hindrar inte att man kan dra slutsatser om skillnaden. Data kommer från ett verkligt datamaterial som följer med R-paketet och beskriver fertiliteten i Schweiz 1888, men resten av uppgifterna i påhittade.

(a) Vad är \bar{x} , \bar{y} , s_x och s_y ? (5p)

(b) Gör lämpliga antaganden om oberoende och gemensam varians och gör ett 95 procents konfidensintervall för skillnaden i fertilitet mellan de två typerna av byar. Kan man säga att det skiljer i fertilitet mellan katolska och icke-katolska byar? (5p)

5. Bonnie Parker har rånat banker i flera år och börjar bli uttråkad. Denna sommar - året innan hon går i kloster, som hon alltid tänkt är det egentliga syftet med hennes liv - bestämmer hon sig för att lära sig lite psykologi medan hon ändå håvar in rånpenningar. Hon bestämmer sig för att undersöka betydelsen av kläder (A) och av språk (B).

Hon testar A på två nivåer: A- är en enkel rånarluva och svart kroppstrumpa. A+ hårda läderkläder med många nitar, som får henne att se ut som en urspårad pundare. B- betyder ett vårdat språk medan B+ är gettoslang med onämnbare svordomar. Hon lägger upp en försöksplan där hon testar alla faktorkombinationer fyra gånger och mäter sin framgång i värdet varje rån ger (Tusentals dollar) i en fullständigt randomiserad design.

A	B					\bar{x}_i	s_i
-	-	8.6	7.3	6.5	8.3	7.67	0.96
+	-	11.5	13.0	13.2	10.80	12.12	1.16
-	+	8.8	5.2	9.00	8.4	7.85	1.78
+	+	12.1	11.6	12.30	10.9	11.72	0.62

(a) Skatta huvudeffekter och samspel samt avgör vilka som är signifikanta. (7p)

(b) Rita en figur som beskriver samspelet och tolka resultatet i a-uppgiften. (3p)

6. Teorifrågor. Ge koncisa svar på nedanstående frågor, inga långa utredningar. Delfrågorna ger 4p vardera.

(a) Det ska vara kalas, och följande händelser definieras:

A : Jimmie kommer

B : Janne kommer

$$\mathbb{P}(A) = \mathbb{P}(B) = 0.5 \text{ och } \mathbb{P}(A \cup B) = 1$$

Vad är $\mathbb{P}(A \cap B)$? Beskriv resultatet med vardagsspråk!

(b) Slumpvariablerna X och Y är oberoende och har samma fördelning $N(\mu, \sigma^2)$, där $\mu = 1$ och $\sigma^2 = 2$. Vilken fördelning (inklusive parametrar) har $-8X$ respektive $X - 2Y$?

(c) En livförsäkring är så uppbyggd att om man dör inom de närmaste 10 åren, så får familjen en 1000 000 kronor. Om man blir invalidiserad, så får man en 500 000. Råkar man ut för en svår olycka, som inte ger bestående men, så får man 250 000. Ersättningen betalas bara ut en gång. Försäkringen betalas med en engångssumma. (I uppgiften ignoreras effekter av inflation.)

Sannolikheterna är:

- död: 1 procent
- invaliditet: 3 procent
- svår olycka utan bestående men: 5 procent.

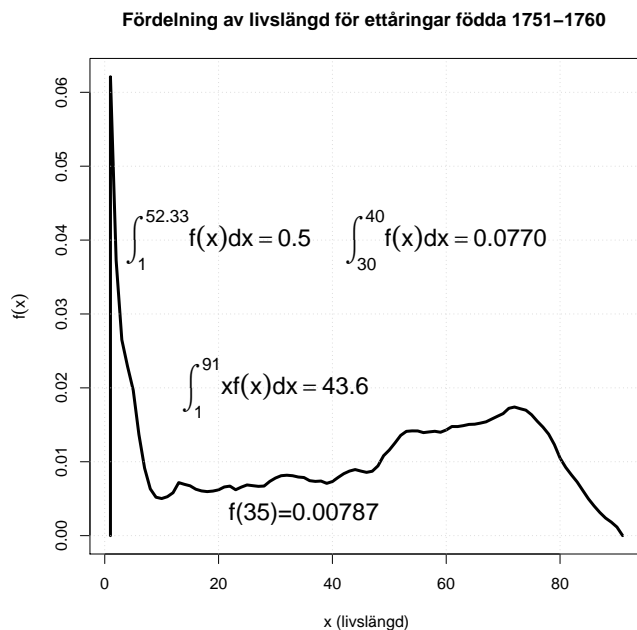
Vad är det minsta försäkringen kan kosta? Dvs. vad är det förväntade värdet av den utbetalning bolaget behöver göra? Att ta mindre betalt för försäkringen än de förväntade utbetalningarna gör inget försäkringsbolag.

(d) $f(x)$, som ritats i figuren nedan, är täthetsfunktionen för

X : Livlängd för en ettårig svensk född 1751 – 1760

För enkelhets skull har antagits att ingen överlevde sin nittioettårsdag (tyvärr ett antagande som inte är långt ifrån verkligheten). Vad är

- Medianen för X ?
- Väntevärdet $\mathbb{E}(X)$?
- $\mathbb{P}(X = 35)$?
- $\mathbb{P}(30 \leq X < 40)$?



(Lägg märke till att X behandlas som en kontinuerlig slumpvariabel, dvs. livslängden mäts inte i hela år utan kan ha decimaler.)

(e) Sonny har tre metoder för att få sina barn att städa rummet på fredagarna:

A: att strängt säga åt dem

B: att hota med att dra in lördagsgodiset (fast han har aldrig satt hotet i verket)

C: att framhålla hur mycket han jobbar för att de ska ha råd att bo kvar i huset sedan mamma gick in i väggen. De kan väl i alla fall hjälpa till!

Han tycker egentligen inte att någon fungerar så bra men han vill i alla fall testa om det finns någon skillnad mellan metoderna. Han kommer på två sätt att göra en försöksplan, där varje kolumn motsvarar en vecka. Se tabellen nedan!

I *Plan 1* har han för varje barn separat slumpmässigt valt ut två veckor då metod A används, två veckor då metod B använts och två veckor då metod C används. I *Plan 2* har alla 18 kombinationer barn-vecka noterats och sedan har 6 av dessa slumpmässigt valts ut för metod A, 6 för metod B och 6 för metod C. Vilken av de två planerna bör han välja? Varför?

	Plan 1						Plan 2					
Vecka	1	2	3	4	5	6	1	2	3	4	5	6
Felicia	B	B	C	A	C	A	C	B	B	A	A	A
Elliot	C	B	A	C	B	A	C	C	B	C	B	A
Tindra	B	B	A	A	C	C	C	A	A	B	B	C

Lycka till!