Matematisk statistik LKT325 (Kemis kurs)
Tentamen 2018-01-09

Tid: 8.30-12.30
Hjälpmedel: **Kursboken Matematisk Statistik av Ulla Dahlbom och Håkan Blomqvists formelsamling.** Boken och formelsamlingen får ej innehålla extra anteckningar, men understrykningar, sticks och markeringar är tillåtna. **Chalmersgodkänd räknare.**
Examinator: Johan Tykesson
Telefonvakt: Johan Tykesson, 0703182096. Rond ca 09.30 och 11.30.
Betygsgränser: för betyg 3 krävs minst 20 poäng, för betyg 4 krävs minst 30 poäng, för betyg 5 krävs minst 40 poäng.

OBS: Till varje uppgift skall fullständig lösning lämnas! Glöm ej att motivera eventuella approximationer. Text på TRE sidor! Lycka till!

1. (3+3 poäng) Antag att ξ är en kontinuerlig stokastisk variabel med frekvensfunktion

$$f(x) = \begin{cases} \frac{3}{2}x - \frac{3}{4}x^2 & \text{för } 0 \leq x \leq 2 \\ 0 & \text{för övrigt.} \end{cases}$$

   (a) Beräkna väntevärde och standardavvikelse för ξ.

   (b) Antag att Annas utgifter under en given dag kan betraktas som en kontinuerlig stokastisk variabel η där $\eta = 700\xi$. Utgifterna under olika dagar kan antas oberoende av varandra. Beräkna (approximativt) sannolikheten att hennes totala utgifter under ett år är mindre än eller lika med 245000 kronor. Antag här att ett år har 365 dagar.

2. (2+2+2 poäng) I en urna ligger 7 gröna bollar och 7 gula bollar blandade. Man drar 5 bollar slumpmässigt, utan återläggning. Låt A vara händelsen att man bara får gröna bollar, och låt B vara händelsen att man får minst en boll av varje färg.

   (a) Beräkna $P(A)$.

   (b) Beräkna $P(A \cup B)$.

   (c) Beräkna $P(A \mid B^c)$.

3. (4 poäng) Man undersöker genomsnittsvikten för vuxna hanhundar av rasen *strävhårig vorsteh*. Man mäter vikterna på fyra slumpmässigt utvalda hundar. Vi antar att vikterna kommer från en normalfördelning med okänt väntevärde μ och okänd standardavvikelse σ, och att vikterna är oberoende. Man får mätvärdena (i kilogram):

$$33.4 \qquad 32.1 \qquad 32.8 \qquad 40.5$$

   Beräkna ett 95% konfidensintervall för μ med hjälp av dessa mätvärden.

4. (7 poäng) En fiskare skall fiska. Han tar sig till en av följande två fiskesjöar: sjö A eller sjö B. Först måste han ta bussen. Tiden han måste vänta på bussen antas rektangelfördelad på intervallet $(0, 12)$, där enheten är minuter. Om han behöver vänta mer än 9 minuter går han till sjö B, annars går han till sjö A. Om han går till sjö A blir antalet fiskar han fångar Poissonfördelat med väntevärde 5 stycken. Om han går till sjö B blir antalet fiskar han fångar Poissonfördelat med väntevärde 4 stycken. Antag att vi efter fisketuren får reda på att han fångade fler än eller lika med 3 fiskar. Beräkna den betingade sannolikheten att han gick till sjö A givet detta.

5. (3+3+1 poäng) Man genomförde ett fullständigt faktorförsök för att undersöka hur de fyra faktorerna A, B, C och D påverkade en viss situation. Man fick följande resultat från de sexton försöken:

Nr.	A	B	C	D	Resultat y
1	-	-	-	-	42
2	+	-	-	-	54
3	-	+	-	-	43
4	+	+	-	-	52
5	-	-	+	-	41
6	+	-	+	-	56
7	-	+	+	-	44
8	+	+	+	-	55
9	-	-	-	+	41
10	+	-	-	+	53
11	-	+	-	+	44
12	+	+	-	+	51
13	-	-	+	+	43
14	+	-	+	+	57
15	-	+	+	+	48
16	+	+	+	+	59

(a) Beräkna huvudeffekten l_A och samspelseffekten l_{AB}.

(b) Antag att man också var intresserad av faktorerna E, F och G. Man gör ett reducerat faktorförsök med 16 försök, där inställningarna för A, B, C och D är som ovan. Man har valt följande generatorer: $E = ABCD$, $F = ABC$ och $G = BCD$. Beräkna alla *ord* (dvs, alla "I") i det reducerade faktorförsöket och bestäm upplösningen. Förklara också varför det reducerade faktorförsöket är dåligt ifall man tycker det är viktigt att E inte blandas ihop med något tvåfaktor samspel.

(c) Antag att man gör ett fullständigt faktorförsök med 8 faktorer A, B, C, D, E, F, G och H. Hur många olika 3-faktor samspel finns det?

6. (4+5 poäng) Antag att man har n glödlampor. Lampornas livslängder antas oberoende och exponentialfördelade med väntevärde 100 timmar. Antag att vi tänder alla lamporna samtidigt och låter dem vara tända tills de går sönder.

   (a) Antag nu att $n = 100$. Beräkna (approximativt) sannolikheten att fler än 40 lampor lyser efter 105 timmar.

   (b) Antag nu att $n = 5$. Låt ξ vara tiden (räknat i timmar från när lamporna tänds) det tar tills två av de fem lamporna gått sönder. (Med andra ord: ξ är tiden det tar tills det återstår tre lysande lampor). Då är ξ en kontinuerlig stokastisk variabel. Beräkna fördelningsfunktionen för ξ.

7. (5 poäng) Zdeno Chara är en av de hårdast skjutande ishockeyspelarna i världen. Inför kommande säsong provar han ut en ny klubba att spela med och han testar bland annat hur hårt han skjuter med varje klubba. Han skjuter elva skott med var och en av de fyra klubborna han testar och skotthastigheterna uppmäts. Här är en påbörjad ANOVA-tabell som tagits fram från mätningarna.

Variationskälla	SS	df	MS	F
Mellan klubbor	184.17	...	...	...
Inom klubbor	983.52	...	...	
Totalt	...	...		

Komplettera ANOVA-tabellen och utför analysen för att undersöka om det är någon skillnad på hur hårt Zdeno skjuter med de olika klubborna i genomsnitt. Tolka även resultatet. Var noga med att berätta vilken hypotes du testar.

8. (4+1+1 poäng) Sven jobbar på ett företag som producerar müsli. En av deras sorter innehåller mycket russin. Det är meningen att varje paket av denna sort ska innehålla ca 200 russin men det har inkommit klagomål på att den innehåller för många russin. Sven får därför i uppgift att testa om müslipaketen i genomsnitt innehåller fler än 200 russin. Om så är fallet måste de göra något åt problemet. Slumpmässigt väljer Sven ut 12 paket och räknar hur många russin det är i varje paket. Medelvärdet av stickprovet beräknades till $\bar{x} = 206.4$ och stickprovsstandardavvikelsen blev $s = 14.5$.

   (a) Hjälp Sven att utföra lämpligt hypotestest. Antag för enkelhetens skull att antalet russin i paketen är normalfördelat. Använde signifikansnivå $\alpha = 0.05$.

   (b) Medelvärdet av stickprovet är större än 200. En av cheferna säger att på grund av detta så kan vi direkt dra slutsatsen att det generellt sätt är för många russin i paketen. Förklara varför det inte är rimligt att resonera på detta sätt.

   (c) Antalet russin i paketen är egentligen inte normalfördelat. Varför inte? (En förklaringar ger ett poäng)