UPPSALA UNIVERSITET
Matematiska institutionen
Rolf Larsson

TENTAMEN I MATEMATISK STATISTIK
Inferens 1 1MS035
2015–12–21

Skrivtid: 8.00-13.00. För betygen 3,4, resp. 5 krävs 18, 25 resp. 32 poäng, inklusive ev. bonuspoäng. Lösningarna ska vara väl motiverade. Skriv endast på ena sidan, börja ny uppgift på ny sida och använd ej rödpenna. Uppgifterna är inte ordnade i svårighetsgrad!

Tillåtna hjälpmedel: Räknedosa. Formelsamling för inferensteori 1.

1. Vi har ett slumpmässigt stickprov $x_1, ..., x_n$ från X, som följer den s.k. lognormalfördelningen, med täthetsfunktion

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-(\ln x - \mu)^2/(2\sigma^2)},$$

för $x \geq 0$ (och 0 annars). Vi antar att parametern σ^2 är känd.
I deluppgifterna nedan får utan bevis användas att $E(X) = e^{\mu+(\sigma^2/2)}$.

   (a) Beräkna momentskattningen av μ.

   (b) Beräkna minsta kvadratskattningen av μ.

   (c) Beräkna maximum likelihoodskattningen av μ.

2. År 2035 genomfördes den första bemannade resan från Jorden till Mars. Väl där gav sig Carl och Carla ut för att mäta tyngdaccelerationen i enheten m/s^2. Oberoende av varandra gjorde Carl fem och Carla fyra sinsemellan oberoende mätningar med resultat enligt tabell nedan.

Carl	3.3	3.8	3.6	3.7	4.1
Carla	3.7	3.7	3.8	3.8	

Antag att Carls mätningar kan ses som ett slumpmässigt stickprov från en slumpvariabel X med väntevärde μ och varians $2\sigma^2$, medan Carlas mätningar kan ses som ett slumpmässigt stickprov från en slumpvariabel Y med väntevärde μ och varians σ^2. Låt $\bar{x}$ och $\bar{y}$ beteckna stickprovsmedelvärdena.

Carl och Carla enades om att kombinera sina skattningar till

$$\mu_1^* = \frac{5\bar{x} + 8\bar{y}}{13}.$$

Resans överbefälhavare Carol tyckte istället att skattningen skulle vara

$$\mu_2^* = \frac{\bar{x} + \bar{y}}{2}.$$

   (a) Beräkna skattningarna numeriskt.

   (b) Visa att båda skattningarna är väntevärdesriktiga.

   (c) Vilken skattning är effektivast?

3. Företaget Plisteriplast tillverkar plastjulgranar i jämn takt. Antalet felaktiga granar kan beskrivas med hjälp av en Poissonprocess med intensitet λ stycken per vecka. Detta innebär (som bekant) att antalet felaktiga granar under t veckor är Poissonfördelat med parameter λt.

   Företagets målsättning är att i medeltal högst en plastjulgran per vecka ska vara felaktig.

   (a) Under första veckans produktion blev det fel på tre granar. Försök att avgöra om företagets målsättning är uppfylld genom att utföra ett lämpligt hypotestest.

   (b) Vilken styrka har testet i (a) om det i själva verket är så att i medeltal tre granar per vecka är felaktiga? Använd en signifikansnivå på högst 5%.

   (c) Överingenjör Granskog är missnöjd med svaret i (b). Han beordrar personalen att utvärdera testet först efter så många veckor att styrkan för ett test med signifikansnivå högst 5% ska vara minst 80% under förutsättning att i medeltal tre granar per vecka är felaktiga. Hur många veckor behövs för detta? Räkna bara med hela veckor.

4. Jesper och Rolf har funnit ett datamaterial på internet som visar antalet dagar med extremt hög nederbörd i Uppsala. Det finns data för åren 1900-2014, där ett år saknas, vilket ger totalt 114 år.

   Jesper och Rolf tänker sig att det de har är ett slumpmässigt stickprov $x_1, ..., x_n$ från en slumpvariabel X med okänt väntevärde μ och okänd varians σ^2. Det visar sig att medelvärdet är $\bar{x} = 4.30$ och stickprovsvariansen är $s^2 = 5.40$.

   (a) Beräkna ett 95%-igt konfidensintervall för μ utan att anta något om fördelningen för X.

   (b) Beräkna ett 95%-igt konfidensintervall för μ under antagandet att X är Poissonfördelad.

   (c) Under antagandet att X är Poissonfördelad, beräkna ett 95%-igt konfidensintervall för sannolikheten att det inte förekommer någon dag med extremt hög nederbörd under ett år.

5. Tioåringarna Johannes och Lina häller knäck i formar. De har 100 formar var. Johannes spiller knäck utanför formen 16 gånger, och Lina gör det 12 gånger.

   (a) Avgör, genom att utföra ett lämpligt hypotestest, om barnen kan anses vara lika duktiga på att hälla knäck i formar.

   (b) Johannes pappa, som är statistiker (och som dessutom gillar knäck), tycker att detta försök var alldeles för litet, och börjar planera för ett större till julen därpå. Om man antar att Johannes och Lina var för sig presterar likadant då som i år, hur många formar var ska de hälla knäck i för att man ska kunna finna en signifikant skillnad mellan deras förmåga på nivån 5%?

6. Enligt Tannenbaums lag är tiden från och med att en julgran tagits in i vardagsrummet tills dess att den har börjat barra exponentialfördelad med väntevärde μ. Julgransförsäljaren Julia vill beräkna ett 99%-igt konfidensintervall för μ. Hon utgår från en undersökning baserad på förra årets kunder. Det ingick 50 kunder i denna undersökning, och de kan anses vara slumpmässigt utvalda. Vid en sammanställning av siffrorna visade det sig att det för dessa kunder i medeltal tog 3.7 dagar innan deras granar började barra.

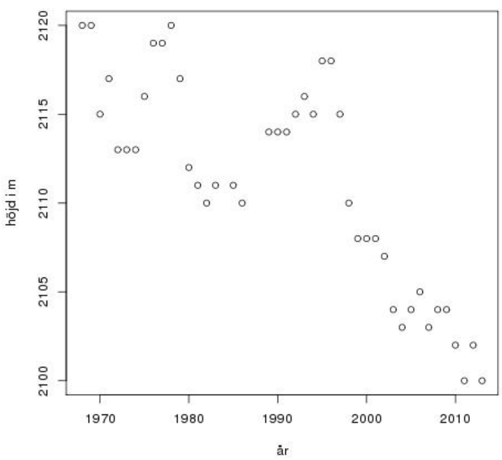
   Hjälp Julia att beräkna konfidensintervallet!

7. Tomtefar tycker att nissarna börjar bli lite för långsamma på att tillverka talande dockor. För att åtgärda detta ber han Tomtemor framställa ett prestationshöjande extrakt på malna renhorn och kottar. För att kontrollera om detta fungerar låter han sedan nissarna på tid tillverka varsin docka, därefter inta extraktet och sedan tillverka en docka till. Försöket utförs med fem slumpmässigt utvalda nissar. Resultaten, i minuter och delar av minuter, ges i tabellen nedan ("tid före" respektive "tid efter" anger respektive tomtes tillverkningstid före och efter att de intagit Tomtemors extrakt). Antag att tillverkningstiderna är normalfördelade, och försök att avgöra om extraktet har avsedd effekt genom att genomföra ett lämpligt hypotestest.

   Var noga med att ange vilka förutsättningar du gör!

Nisse nr	1	2	3	4	5
Tid före	7.3	6.4	11.8	2.7	8.9
Tid efter	5.2	4.1	8.7	3.2	6.5

8. Höjden på Kebnekajses sydtopp i meter bokfördes under åren 1968-2013, med några års bortfall. Totalt omfattar denna serie 43 år. Mätningarna visas i figuren nedan.



Antag att höjden på Kebnekajses sydtopp Y_t år t följer regressionsmodellen

$$Y_t = \alpha + \beta t + \varepsilon_t$$

där alla ε_t är oberoende $N(0, \sigma^2)$.

Räknehjälp: $\bar{t} = 1990.79$, $\bar{y} = 2111.12$,

$$\sum_{t=1}^{43}(t-\bar{t})^2 = 8043.12, \quad \sum_{t=1}^{43}(y_t-\bar{y})^2 = 1504.42, \quad \sum_{t=1}^{43}(t-\bar{t})(y_t-\bar{y}) = -2812.95.$$

(a) Skatta parametrarna α och β.

(b) Beräkna förklaringsgraden.

(c) Beräkna ett 95%-igt konfidensintervall för β.

(d) Uppskatta vilket år som höjden för första gången kommer att understiga 2090 meter.

LYCKA TILL!