TMS051 Matematisk statistik och simuleringsteknik, Z, del A

Tentamen 31 augusti 2006 fm

Även omtentamen på TMS050 Matematisk statistik och simuleringsteknik Z, del A.

Tillåtna hjälpmedel är räknedosa utan information om kursen i minnena, Beta samt kursens formel- och tabellsamling.

Examinator är Tommy Norberg, ankn 3528 eller 0730 79 42 09. Tommy går att nås per mobiltelefon under tentamen.

Maximalt antal tentamenspoäng är 30, av dessa krävs 12 för godkänt betyg på tentan och 18 resp 24 för 4:a och 5:a. Lösningar går att ladda ner från kurshemsidan. Rättningsprotokoll anslås i MV:F, plan 2.

Svar skall motiveras om ej annat sägs i uppgiften.

Uppgifter

1. Varje dag skippas 20 enheter till en sammansättningsindustri. Fyra av dessa sorteras ut slumpmässigt och funktionskontrolleras. Antag att sex enheter ej fungerar tillfredsställande. Hur stor är sannolikheten att man hittar

   (a) exakt två enheter som ej fungerar tillfredsställande? (2 p)

   (b) minst två enheter som ej fungerar tillfredsställande? (2 p)

2. När Kvast-Hilda, stadens mest kända häxa, kokar ihop en soppa som kan förvandla en groda till en prinsessa händer det ibland att hon istället får en soppa som istället förvandlar grodan till en häst. Vis av mångårig erfarenhet vet Kvast-Hilda att detta händer ca 3 gånger av 17. Under tre på varandra följande veckor kokar Kvast-Hilda denna soppa. Ungefär hur stor är sannolikheten att det blir fel

   (a) en gång? (2 p)

   (b) minst en gång? (2 p)

3. Beräkna väntevärde och 75%-kvantil för exponentialfördelningen. (3 p)

4. Antag att tiden tills en viss utrustning "går ner" (alltså slutar att fungera) är exponentialfördelad med väntevärde μ. Parametern $\lambda = 1/\mu$ brukar då kallas för intensiteten. I ett test av 7 sådana utrustningar uppmättes funktionstiderna

   76.6 4.3 3.8 1.7 30.1 4.2 22.5

   ML-skatta intensiteten λ. (4 p)

5. Man mätte tiden t för utförandet av en viss uppgift i sammansättningen av en viss produkt. I $n = 100$ på varandra följande utföranden erhölls $\sum_i t_i = 7436$ s, $\sum_i t_i^2 = 553948$. Under lämpliga villkor kan man punkt- och intervallskatta väntevärdet μ. (a) Vilka är dessa villkor? (b) Gör nu punkt- och intervallskattningen. Den senare ska ha konfidensen 99%. (c) Granska kritiskt villkoren. Finns det något — och i så fall, vilket — som kanske inte är uppfyllt? (1+3+1 p)

6. Under samma villkor som i ovanstående problem kan man intervallskatta standardavvikelsen σ. Intervallet ska vara uppåt begränsat (men ej nedåt) och konfidensen ska här vara 99%. (3 p)

7. Nu när vi börjar närma oss valtider görs många politiska partundersökningar. I en fick alliansen 50.1% medan det andra blocket tillsammans erhöll 45.2%. Antag att 2 000 personer besvarade enkäten och att den är gjord enligt vedertagen statistisk metodik. Intervallskatta med konfidensen 95% skattningen av sannolikheten att en godtycklig väljare röstar på något av allianspartierna. (3 p)

8. Antar man att inga nya partier kommer in i riksdagen, så är det bara $(0.501 + 0.452) \cdot 2\,000 = 1\,906$ av de 2 000 avlagda svaren som bör räknas, eftersom det är rösterna på partierna som kommer in i riksdagen som avgör mandatfördelningen. Är alliansens ledning med detta sätt att räkna statistiskt säkerställd? (4 p)

Lycka till!

Lösningar till TMS051, del A den 31/8-06

1. (a) $p = \binom{6}{2}\binom{14}{2}/\binom{20}{4} = 0.282$

   (b) $p = 1 - \left(\binom{6}{0}\binom{14}{4}/\binom{20}{4} + \binom{6}{1}\binom{14}{3}/\binom{20}{4}\right) = 0.343$

2. (a) $p = \binom{3}{1}\left(\frac{3}{17}\right)^1\left(\frac{14}{17}\right)^2 = 0.359$

   (b) $p = 1 - \binom{3}{0}\left(\frac{3}{17}\right)^0\left(\frac{14}{17}\right)^3 = 0.441$

3. Exponentialfördelningens täthet är $f(x) = \lambda e^{-\lambda x}$ eller $f(x) = \frac{1}{\beta}e^{-x/\beta}$, beroende på om man väljer intensiteten λ eller $\beta = 1/\lambda$ som parameter. Utfallsrummet är positiva talaxeln, så $x > 0$, Med den förstnämnda parametriseringen fås

$$\mu = \int_0^\infty x\,\lambda e^{-\lambda x}\,dx = \{\text{part int}\} = 1/\lambda$$

och

$$0.75 = \int_0^{x_{0.75}} \lambda e^{-\lambda x}\,dx = 1 - e^{-\lambda x_{0.75}}$$

$$\Rightarrow \quad 0.25 = e^{-\lambda x_{0.75}} \quad \Rightarrow \quad x_{0.75} = \frac{-\ln 0.25}{\lambda} = \frac{\ln 4}{\lambda}$$

4. Tätheten, uttryckt m.h.a parametern λ, är $f(x) = \lambda e^{-\lambda x}$, $x > 0$. Utfallet $x_1, \ldots, x_n$ har därför troligheten

$$L(\lambda) = f(x_1)\cdot \ldots \cdot f(x_n) = \lambda e^{-\lambda x_1}\cdot \ldots \cdot \lambda e^{-\lambda x_n} = \lambda^n e^{-\lambda \sum_i x_i}$$

ML-skattningen är det λ som maximerar trolighetsfunktionen. Logaritmera, sätt derivatan till noll och lös ut λ:

$$\mathcal{L}(\lambda) = \log L(\lambda) = n\log\lambda - \lambda\sum_i x_i$$

$$\mathcal{L}'(\lambda) = \frac{n}{\lambda} - \sum_i x_i$$

$$\mathcal{L}'(\lambda) = 0 \quad \Leftrightarrow \quad \frac{n}{\lambda} = \sum_i x_i \quad \Leftrightarrow \quad \lambda = \frac{n}{\sum_i x_i} = \frac{1}{\bar{x}}$$

ML-skattningen är således $\hat{\lambda} = 1/\bar{x}$. Till sist, $\sum_i x_i = 76.6 + 4.3 + 3.8 + 1.7 + 30.1 + 4.2 + 22.5 = 143.2 \Rightarrow \hat{\lambda} = 7/143.2 = 1/20.457 = 0.0489$.

5. (a) Observationerna måste vara oberoende och likafördelade

   (b) Vi får $\bar{t} = 7436/100 = 74.36$ och $s^2 = \frac{1}{99}\left(553948 - \frac{1}{100}7436^2\right) = 10.1721 \Rightarrow s = 3.189$ etc. Vi får nu att $\mu = \bar{t} \pm 2.6s/\sqrt{n} = 74.4 \pm 0.83$ etc. När man har så många observationer som här är det ingen mening med att hämta kvantilen ur t-tabellen. Det är tillräckligt noggrant att ange det statistiska 99%-felet till 2.6 gånger standardfelet. Hade önskad konfidensgrad varit 95%, hade jag istället multiplicerat standardfelet med 2.

   (c) Oberoendeantagandet behöver verifieras på något sätt, eftersom det typiskt är samma operatör som gör de 100 monteringarna och det kan ju finnas beroenden mellan på varandra följande monteringar. Dessutom skulle jag vilja veta säkert att man inte bytt operatör under försöket, för då kanske inte observationerna är likafördelade.

6. Vi utgår ifrån att $(n-1)s^2/\sigma^2 \sim \chi^2(n-1)$. I $\chi^2(100)$-tabellen hittar vi 0.99-kvantilen 70.0649, vilket värde är aningen för stort, eftersom antalet frihetsgrader är 99 inte 100. Händelsen $70.0649 \leq 99s^2/\sigma^2$ inträffar alltså med den ungefärliga sannolikheten 0.99, och genom att stoppa in $s^2 = 10.172$ och omvandla olikheten får vi att det sökta 99%-intervallet är $\sigma \leq \sqrt{99 \cdot 10.172/70.0649} = 3.79$. Se fotnot 1.

7. $p = 0.501 \pm 2\sqrt{0.501 \cdot 0.499/2000} = 0.501 \pm 0.022$ ("felet" är alltså drygt 2 procentenheter, trots att man tillfrågar så många som 2000 väljare)

8. Av 1906 tillfrågade väljare är det $0.501 \cdot 2000 = 1002$ som väljer något alliansparti. Ett nedåt begr konfidensintervall för proportionen alliansväljare är därför $p \geq 1002/1906 - 1.65\sqrt{\frac{1002}{1906}\left(1 - \frac{1002}{1906}\right)/1906} = 0.526 - 0.019 = 0.507$. Det förefaller alltså som att alliansen hade ett övertag då mätningen gjordes. (Obs 1.65 är 0.95-kvantilen i normalfördelningen.)

Fotnot 1 I ett tidigare förslag till lösning av uppgift 6 användes felaktigt 0.01-kvantilen 135.807. Tack till er som påpekade detta misstag. Till alla: att $\sigma \leq 2.7$ gäller med 99% konfidens, vilket ju var svaret i den felaktiga lösningen, är orimligt eftersom $s = 3.189$. Däremot gäller $\sigma \geq 2.7$ med 99% konfidens, men det är en annan historia.

Resultat:

	U	3	4	5	S:a
antal	0	3	0	0	3
%					

TMS051 Matematisk statistik och simuleringsteknik, Z, del A

Tentamen 22 oktober 2005 em V

Även omtentamen på TMS050 Matematisk statistik och simuleringsteknik Z, del A.

Tillåtna hjälpmedel är räknedosa utan information om kursen i minnena, Beta samt kursens formel- och tabellsamling.

Examinator är Tommy Norberg, ankn 3528 eller 0730 79 42 09.

Maximalt antal tentamenspoäng är 30, av dessa krävs 12 för godkänt betyg på tentan och 18 resp 24 för 4:a och 5:a. Lösningar går att ladda ner från kurshemsidan. Rättningsprotokoll anslås i första våningen i Matematisk centrum, Eklandagatan 86. Genomgång av tentan görs torsdagen den 10 november kl 13-15 i HA3. Granskning av tentan kan under terminstid även göras i mottagningsrummet på entréplanet (våning 2) i Matematisk centrum må-fr 12^{30}–13. Obs att Matematiskt centrum flyttar under juluppehållet.

Svar skall motiveras om ej annat sägs i uppgiften.

Uppgifter

1. Visa för två godtyckliga händelser A och B, att

   (a) $P(B \setminus A) = P(B) - P(A \cap B)$ (2 p)

   (b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (2 p)

2. Du drar två kort ur en välblandad kortlek. Beräkna sannolikheten att båda har samma valör. (För den som aldrig spelat kort eller glömt: En kortlek består av 52 kort, indelade i 13 valörer med 4 kort i varje.) (3 p)

3. I syfte att ta reda på om det verkligen är någon skillnad i smak på en viss känd läskedryck och en billigare kopia gjordes 10 s.k dubbla blindtest, i vilka en försöksperson ombedes tala om vilken av två drycker han/hon tycker bäst om. (Att testet är ett dubbelt blindtest betyder att inte heller försöksledaren vet vilken av dryckerna som är originalet.) Låt X vara antalet gånger försökspersonen väljer originalet. (a) Vilken sannolikhetsfördelning har X om det inte är någon skillnad på dryckerna? I 8 av de 10 försöken valde försökspersonen originalet. (b) Är ett så pass extremt resultat förvånande om det inte är någon skillnad på dryckerna? Tänk dig in i situationen att du är försöksledare och vill dra någon form av slutsats av experimentets resultat. Svara kortfattat, men tydligt. (2+2 p)

4. För en produktionsanläggning vet man att den går ned med intensiteten ≈ 0.0025 per timma. Beräkna under lämpligt antagande (a) förväntat antalet stopp under en period av 400 timmar, och (b) förväntad tid tills nästa stopp. Glöm inte att tydligt ange vilket antagande du gör. (2+2+1 p)

5. Antag att de två stokastiska variablerna X och Y är kontinuerligt fördelade med täthet

   $$f(x, y) = c \quad \text{för} \quad 0 \le x \le y \le 1$$

   Beräkna $\mathrm{Cov}[X, Y]$. (4 p)

6. Besvara med Ja eller Nej följande tre frågor. Motivera inte dina svar.

   (a) Är medelvärdet en väntevärdesriktig skattning av väntevärdet? (1 p)

   (b) Är stickprovsvariansen en väntevärdesriktig skattning av variansen? (1 p)

   (c) Är stickprovets standardavvikelse en väntevärdesriktig skattning av standardavvikelsen? (1 p)

7. Under fem konsekutiva arbetsdagar lät man fem medarbetare utföra ett visst moment i produktionen av en viss vara. Efter introduktion och en timmas träning under överinseende av ansvarig arbetsledare, mätte man tiden det tog att utföra momentet. Härvid erhölls för de fem medarbetarna:

$$9.13 \quad 6.67 \quad 10.25 \quad 10.58 \quad 7.71$$

   (tidsenhet: minuter) Syftet var att skaffa sig en övre gräns för momentets snittid. Den på företaget ansvarige för statistisk bearbetning konstaterade efter att ha gjort en QQ-plot[1] att det inte är orimligt att i den fortsatta analysen utgå ifrån att tiderna är (i alla fall approximativt) normalfördelade. Låt μ beteckna den förväntade tiden det tar att utföra momentet. Beräkna en övre gräns μ_0, sådan att risken att påståendet $\mu \leq \mu_0$ är fel är ca 5%. (4 p)

8. I syfte att få en uppfattning av kvaliteten hos en viss vara, tillverkad av en underleverantör, valde man i 100 konsekutiva skeppade partier slumpmässigt ut en enhet ur varje. Den utvalda enheten kontrollerades noggrant. Här gör vi det förenklande antagandet att kontrollen var m.a.p en viss funktionalitet och att man helt enkelt avgjorde om enheten höll de med underleverantören överenskomna kraven för denna funktionalitet eller ej. Det visade sig att 90 uppfyllde kraven och att 10 ej gjorde det. Punkt- och intervallskatta sannolikheten att en godtycklig enhet levererad av denna underleverantör uppfyller kraven för varan ifråga. (3 p)

Lycka till!

[1] I del B får du lära dig vad en Q(uantile-)Q(uantile)-plot är och vad man kan använda en sådan till

Lösningar till TMS051, del A den 22/10-05

1. (a) $P(B) = P(A \cap B \cup B \setminus A) = P(A \cap B) + P(B \setminus A)$

   (b) $P(A \cup B) = P(A \cup B \setminus A) = P(A) + P(B \setminus A) = P(A) + P(B) - P(A \cap B)$

   Här ska man tydligt visa att argumenteringen baseras på axiomen. Annars blir det inga poäng.

2. Totala antalet utfall är $\binom{52}{2} = \frac{52 \cdot 51}{2}$. Antalet utfall där båda korten har samma valör är enligt multiplikationsprincipen $13 \cdot \binom{4}{2} \cdot \binom{48}{0} = 13 \cdot \frac{4 \cdot 3}{2} \cdot 1$. Sannolikheten för lika valör (par) blir således enl den klassiska sannolikhetsdefinitionen

$$\frac{13 \cdot \binom{4}{2} \cdot \binom{48}{0}}{\binom{52}{2}} = 13 \cdot \frac{4}{52}\frac{3}{51} = \frac{3}{51} = \frac{1}{17} \approx 0.0589$$

   Man kan också tänka så här: Första kortet spelar ingen roll. Det bestämmer dock i vilken valör man kan få sitt par. Sedan gäller det att andra kortet får samma valör som det första. Denna sannolikhet är naturligtvis 3/51.

3. (a) Man gör 10 oberoende och likafördelade försök. Om det inte är någon skillnad mellan dryckerna blir valet av dryck fullständigt slumpmässig. Sannolikheten att försökspersonen väljer originalet måste då vara 1/2. Härur drar vi slutsatsen att X är binomialfördelad med parametrar $n = 10$ och $p = 1/2$. (b) Vi har att

$$P(X \geq 8) = \sum_{k=8}^{10} \binom{10}{k} (1/2)^k (1/2)^{10-k} = \frac{45 + 10 + 1}{1024} = \frac{56}{1024} \approx 0.055$$

   Att en försöksperson väljer någon av dryckerna minst 8 gånger av 10, sker alltså oftare än 1 gång på 10 (om det inte är någon skillnad). Jag skulle därför inte bli sådär väldigt förvånad. Resultatet 8-av-10 kan dock tyda på att försökspersonen tycker att originalet smakar bättre. Men att statistiskt säkerställa detta kräver fler försök.

4. Här är det rimligt att antaga att stoppen följer en Poissonprocess, vilket innebär (1) att antalet stopp under tiden $[0, t]$ är $\mathrm{Poi}(\lambda t)$ och (2) att tiden till nästa stopp är $\exp(\lambda)$. Här är $\lambda \approx 0.0025 = 1/400$. Ur formelsamling eller Beta får vi nu reda på att svaret på (a) är $\lambda t = 1$ och att svaret på (b) är $1/\lambda = 400$ timmar.

5. Arean av utfallsrummet för paret (X, Y) är 1/2, så $c = 2$. De två marginaltätheterna för X och Y är

$$f(x) = \int_x^1 c\, dy = 2(1 - x) \quad \text{för} \quad 0 \leq x \leq 1$$
$$f(y) = \int_0^y c\, dx = 2y \quad \text{för} \quad 0 \leq y \leq 1$$

   Vi får nu att

$$E[X] = \int_0^1 2x(1 - x)\, dx = \int_0^1 (2x - 2x^2)\, dx = 1 - \frac{2}{3} = \frac{1}{3}$$
$$E[Y] = \int_0^1 2y^2\, dy = \frac{2}{3}$$
$$E[XY] = \int_0^1 \int_0^y xyc\, dxdy = 2\int_0^1 y \int_0^y x\, dxdy = \int_0^1 y^3\, dy = \frac{1}{4}$$

(Att $E[X] = 1 - E[Y]$ är väl ganska självklart, eller... Titta på marginaltätheterna.) Det följer nu att

$$\mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y] = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}$$

6. (a) Ja, ty $E[\bar{X}] = \mu$

   (b) Ja, ty $E[S^2] = \sigma^2$

   (c) Nej, ty $E[S] \neq \sigma$ utom i ointressanta undantagsfall

7. Vi noterar att $n = 5$, $\sum x = 44.34$, $\sum x^2 = 404.2888$. Härur fås att

$$\bar{x} = \frac{44.34}{5} = 8.868 \quad \text{och} \quad s = \sqrt{\frac{404.2888 - \frac{44.34^2}{5}}{4}} = 1.664$$

Ur t-tabell (4 frihetsgrader) fås kritiska värdet $t_{0.05} = 2.132$ och vi ser till sist att

$$\mu \leq \bar{x} + t_{0.05} s/\sqrt{n} = 8.868 + 1.587 \approx 10.45 \text{ minuter}$$

gäller med konfidensen ca 95%.

8. Punktskattningen av den sökta sannolikheten är $\widehat{p} = 90/100 = 0.90$. Härur fås att $n \min(p, 1-p) \approx 10 > 5$. Vi kan därför använda normalapproximationen av $\mathrm{bin}(n,p)$ då vi beräknar ett konfidensintervall för p. Då inget sagts om konfidensgraden väljer vi att sikta på ≈ 0.95. Motsv kvantil (d.v.s 0.975-kvantilen) i normalfördelningen är ≈ 2, så

$$p = \widehat{p} \pm 2\sqrt{\widehat{p}(1-\widehat{p})/n} = 0.90 \pm 0.06$$

eller

$$p \in [0.84,\ 0.96]$$

gäller med ca 95% konfidens. Med ca 90% konfidens gäller $p = 0.90 \pm 0.05$ och med ca 99% konfidens gäller $p = 0.90 \pm 0.08$.

TMS051 Matematisk statistik och simuleringsteknik, Z, del A

Tentamen 12 januari 2006 fm V

Även omtentamen på TMS050 Matematisk statistik och simuleringsteknik Z, del A.

Tillåtna hjälpmedel är räknedosa utan information om kursen i minnena, Beta samt kursens formel- och tabellsamling.

Examinator är Tommy Norberg, ankn 3528 eller 0730 79 42 09.

Maximalt antal tentamenspoäng är 30, av dessa krävs 12 för godkänt betyg på tentan och 18 resp 24 för 4:a och 5:a. Lösningar går att ladda ner från kurshemsidan. Rättningsprotokoll anslås i nya MV-huset, Hörsalsvägen 1.

Svar skall motiveras om ej annat sägs i uppgiften.

Uppgifter

1. Visa att

$$P(A|B) = P(A) \quad \Leftrightarrow \quad P(B|A) = P(B) \qquad \text{(3 p)}$$

2. För en brusig binär kommunikationslänk gäller att en sänd etta felaktigt tas emot som nolla i ca 5% av fallen och att en sänd nolla felaktigt tas emot som etta i ca 2.5% av fallen. Man räknar med att proportionen ettor som sänds är ca 50%. Antag att en nolla togs emot. Hur stor är den betingade sannolikheten att det var en nolla som sändes? (4 p)

3. Antag att du har för avsikt att göra 5 oberoende upprepningar av ett försök sådant att det lyckas med sannolikheten 0.25. Beräkna sannolikheten att
   (a) minst ett försök lyckas (2 p)
   (b) exakt två försök lyckas (2 p)

4. En stokastisk variabel är normalfördelad med väntevärde 4 och standardavvikelse 0.8. Man har för avsikt att göra fyra oberoende observationer av den. Beräkna sannolikheten att
   (a) den andra observationen hamnar i intervallet från 3.6 till 4.4 (2 p)
   (b) medelvärdet hamnar i intervallet från 3.6 till 4.4 (2 p)

5. Två stokastiska variabler X och Y har standardavvikelserna 3 resp 4. Deras korrelation är -0.875. Beräkna standardavvikelsen för $X + Y$. (3 p)

6. Motivera inte dina svar på denna uppgift.
   (a) Du läser i din dagstidning om den senast utförda partisympatiundersökningen. I texten läser du att Junilistan, om det vore val idag, skulle få 3.7% av rösterna. I den förra undersökningen (som gjordes en månad tidigare) var proportionen väljare som skulle rösta på Junilistan 4.8%. Du får också reda på att förändringen är statistiskt säkerställd. Betyder detta att risken är ca 5% att påståendet

(1) att Junilistan inte skulle klara 4%-gränsen om det vore val idag

(x) att färre än 4.8% procent av väljarna skulle rösta på Junilistan om det vore val idag

(2) att färre skulle rösta på Junilistan idag än fallet var för en månad sedan

är fel? Vilket alternativ är korrekt—1, x eller 2? (2 p)

(b) I en rapport av en statistisk analys får du reda på att ett visst estimat är väntevärdesriktigt. Betyder detta att

(1) om experimentet upprepas många gånger och man beräknar medelvärdet av estimaten, så kommer detta medelvärde att konvergera mot det sanna parametervärdet då antalet upprepningar går mot oändligheten

(2) sannolikheten att det sanna värdet är större resp mindre än det estimerade är 50%?

Vilket alternativ är korrekt—1 eller 2? (2 p)

7. I en studie av ett materials hållfasthetsegenskaper uppmättes följande värden för 5 oberoende prover:

$$2.10 \quad 3.52 \quad 5.14 \quad 3.98 \quad 6.47$$

Enheten är oväsentlig. Antag att observationerna är lognormalfördelade. Låt μ resp σ beteckna väntevärdet och standardavvikelse på log-skalan. Kan man med någon rimlig säkerhet påstå att $e^{\mu} > 2.5$?

8. 20 av ett känt stort bilföretags testförare provkörde en ny testbana så fort de kunde. Varje förare körde 5 varv och den snabbaste varvtiden registrerades. I syfte att undersöka variationen i snabbhet mellan förarna räknades därefter standardavvikelsen ut. Det visade sig att denna blev 8.76 sekunder. Bestäm ett konfidensintervall för motsv teoretiska storhet. Önskad konfidensgrad är ca 95%. (En kvantilplott gjordes. Den sa inte emot antagandet att snabbaste varvtiden är $N(\mu, \sigma)$-fördelad.)

Lycka till!

Svar eller lösningar till TMS051, del A den 12/01–06

1. Lättast är att konstatera att båda påståendena är ekvivalenta med att A och B är oberoende. Då måste de självklart vara ekvivalenta. Men man kan också antaga att $P(A|B) = P(A)$. Då följer att $P(A \cap B) = P(A|B)P(B) = P(A)P(B)$ Härur följer nu att $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B)$. Detta bevisar den ena implikationen. Den andra följer av symmetrin. (Har man visat att $P(A|B) = P(A) \Rightarrow P(B|A) = P(B)$, så har man naturligtvis även visat att $P(B|A) = P(B) \Rightarrow P(A|B) = P(A)$. Det är ju bara att byta ut A mot B och B mot A i den visade implikationen.)

2. Låt S och M beteckna det som sänds resp det som tas emot. Vi får reda på att $P(M = 0|S = 1) \approx 0.05$, att $P(M = 1|S = 0) \approx 0.025$ och att $P(S = 1) \approx P(S = 0) \approx 0.50$. Bayes formel ger

$$P(S=0|M=0) = \frac{P(S=0)P(M=0|S=0)}{P(S=0)P(M=0|S=0) + P(S=1)P(M=0|S=1)}$$
$$\approx \frac{0.50 \cdot (1-0.025)}{0.50 \cdot (1-0.025) + 0.50 \cdot 0.05} = \frac{0.4875}{0.5125} \approx 0.951$$

3. Antalet lyckade försök är Bin$(5, 0.25)$, så (a) $1 - 0.75^5 \approx 0.763$ (b) $\binom{5}{2} 0.25^2\, 0.75^3 \approx 0.264$

4. (a) $\Phi(0.5) - \Phi(-0.5) = 2\Phi(0.5) - 1 \approx 0.383$ (b) $\Phi(1) - \Phi(-1) = 2\Phi(1) - 1 \approx 0.683$

5. Notera först att $\mathrm{Kov}[X, Y] = -0.875 \cdot 3 \cdot 4 = -10.5$. Vi får sedan att $\mathrm{Var}[X + Y] = 3^2 + 4^2 - 2 \cdot 10.5 = 4$. Således har summan standardavvikelsen 2.

6. (a) 2 (b) 1

7. Efter logaritmering av data får jag $\sum x = 6.8859$ och $\sum x^2 = 10.2084$. Dessutom är $n = 5$. Således är $\bar{x} = 1.377$ och $s = \sqrt{0.18132} = 0.4258$. Gränsen i ett nedåt begränsat konfidensintervall för μ är $\bar{x} - t_{0.05,4} s/\sqrt{n} \approx 0.971$ (konfidensgrad ca 95%) och $\mu \geq 0.971$ $\Leftrightarrow$ $e^\mu \geq e^{0.971} = 2.64 > 2.5$. Således gäller att risken att påståendet $e^\mu > 2.5$ är fel är mindre än ca 5%. Svaret på frågan är alltså ja, med iallafall ca 95% konfidens (eller ej mer än ca 5% felrisk).

8. Ur tabell i t.ex formelsamlingen fås kvantilerna $\chi^2_{19,0.025} = 8.9065$ och $\chi^2_{19,0.975} = 32.8523$. Via invertering av den dubbla olikheten

$$\chi^2_{19,0.025} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{19,0.975}$$

(som ju inträffar med sannolikheten 0.95) erhålles

$$6.66^2 = 44.381 = \frac{(n-1)S^2}{\chi^2_{19,0.975}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{19,0.025}} = 163.702 = 12.79^2$$

Det sökta konfidensintervallet är således $6.66 \leq \sigma \leq 12.79$.