F14 – Repetition

Måns Thulin

Uppsala universitet
thulin@math.uu.se

Statistik för ingenjörer • 6/3 2013

Dagens föreläsning

- Tentamensinformation
- Exempel på tentaproblem

På kurshemsidan finns sex gamla exempeltentor; från 2010, 2011 och 2012; med lösningsförslag.

Tentamen 13/3

- Onsdagen 13/3 kl 08.00-13.00 i Polacksbackens skrivsal.
- Tillåtna hjälpmedel:
  - Räknedosa.
  - Tabellblad (motsvarande s. 129-131 i boken. Utdelas på tentan!)
  - Egen formelsamling: ett A4-papper med egna **handskrivna** anteckningar. Man får skriva på **båda sidorna** av pappret. Man får skriva **vad som helst**, exempelvis formler och lösningar på gamla tentaproblem.
- Totalt 40 poäng. A-delen består av 10 poäng och B-delen av 30 poäng.
  - På A-delen krävs minst 8 av 10 poäng för godkänt.
  - Bonuspoäng från diskussions- och inlämningsuppgifter räknas bara till B-delen och endast vid det ordinare tentamenstillfället.
  - Betygsgränser: sammanlagt (A+B) minst 18 för betyg 3, minst 25 för betyg 4 och minst 32 för betyg 5.

A-delen

A-delen på tentan är värd 10 p. Av dessa så måste man få minst 8 p för att kunna bli godkänd.

Bonuspoängen från inlämningsuppgifterna räknas *inte* mot 8 p-gränsen.

Typiska A-delsproblem:

(a) Räkna ut en sannolikhet; typ $\mathrm{P}(A \cap B)$ eller $\mathrm{P}(C|D)$.
(b) Givet sannolikhetsfunktion eller täthetsfunktion, beräkna en sannolikhet.
(c) Tolka lådagram och/eller histogram.
(d) $X \sim N(\mu, \sigma^2)$. Beräkna en sannolikhet för X.
(e) X är *Bin*- eller *Po*-fördelad. Beräkna en sannolikhet för X.
(f) Använd centrala gränsvärdessatsen. (värt 2 p)
(g) Beräkna en konfidensintervall för μ givet $\bar{x}$ och σ^2 eller s^2. (värt 2 p)
(h) Gör ett steg i ett regressionsproblem (en skattning, en prediktion...).

Kursmål

För godkänt betyg ska studenten kunna...

- genomföra enkla beräkningar av sannolikheter;
- redogöra för begreppet stokastisk variabel *(slumpvariabel)* och kunna använda några vanliga sannolikhetsfördelningar;
- tolka centrala gränsvärdessatsen;
- använda punkt- och intervallskattningar för några statistiska typproblem;
- tillämpa enkel regressionsmetodik för anpassning av mätdata;
- ange ett flertal metoder och tekniker för visualisering av datamaterial;
- redogöra för några typiska ingenjörstillämpningar av sannolikhet och statistik, exempelvis tillförlitlighet och kvalitetsteknik.

A-delen från tentan 2012-08-24

(a) Händelserna A och B är oberoende, med $\mathrm{P}(A) = 0.8$ och $\mathrm{P}(B) = 0.2$. Beräkna $\mathrm{P}(A \cup B)$. (1p)

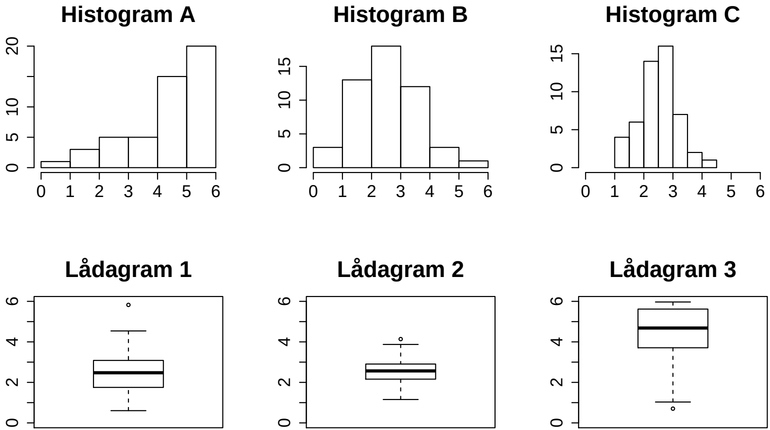
(b) En slumpvariabel X har täthetsfunktionen

$$f(x) = 4x^3, \qquad 0 \leq x \leq 1.$$

Beräkna $\mathrm{P}(X \geq 0.5)$. (1p)

A-delen från tentan 2012-08-24

(c) Nedan har tre datamaterial illustrerats dels med histogram och dels med lådagram. Para ihop rätt histogram med rätt lådagram. (1p)



A-delen från tentan 2012-08-24

(d) En slumpvariabel X är fördelad enligt $X \sim \mathrm{N}(4, 4)$. Beräkna $\mathrm{P}(X \leq 5)$. (1p)

(e) Låt $X \sim \mathrm{Bin}(10, 0.1)$. Beräkna $\mathrm{P}(X \leq 2)$. (1p)

(f) De oberoende slumpvariablerna $X_1, \ldots, X_{2000}$ kommer från en fördelning med $\mathrm{E}[X] = 5$ och $\mathrm{V}[X] = 25$. Ange den approximativa fördelningen för slumpvariabeln $Y = X_1 + \cdots + X_{2000}$. (2p)

A-delen från tentan 2012-08-24

(g) För 15 observationer från $\mathrm{N}(\mu, \sigma^2)$ har man funnit medelvärdet $\bar{x} = -3.14$ och stickprovsvariansen $s^2 = 2.72$. Beräkna ett 99% konfidensintervall för μ. (2p)

(h) För $n = 6$ observationer av variabelparet (x, y) har man fått $\bar{x} = 3.5$, $\bar{y} = 7.6$, $S_{xx} = 17.5$, $S_{yy} = 35.38$ och $S_{xy} = -22.11$. Man antar ett linjärt samband på formen $y = kx + mx$. Skatta k. (1p)

B-delen

B-delen innehåller "läsetal" som för det mesta löses i flera steg. Problemen är sammanlagt värda 30 poäng och bonuspoängen från inlämings- och diskussionsuppgifterna räknas in här.

Typiska B-delsproblem:

(i) Beräkna en sannolikhet eller ett väntevärde givet en täthetsfunktion.

(ii) Räkna ut en sannolikhet med hjälp av centrala gränsvärdessatsen.

(iii) Känna igen binomialfördelningen, beräkna konfidensintervall för parametern p.

(iv) Jämföra två datamaterial med hjälp av konfidensintervall för $\mu_1 - \mu_2$.

(v) Göra olika beräkningar för regressionsproblem (skattningar, konfidensintervall) eller tolka regressionsdata.

(vi) Mindre typiska problem (som dock kan dyka upp): konfidensintervall för μ, konfidensintervall för $p_1 - p_2$, mer avancerade sannolikheter för normalfördelningen (typ hissproblemet), väntevärdesriktighet/varians för skattningar, väntevärde för olika strategier (kretskortproblemet), kluringar...

B-delen från tentan 2012-08-24

2. Vissa exemplar av den nya mobiltelefonen iDroid visar sig ha ett antennfel som gör att de får dålig täckning i 4G-nätet. Bland de 1000 första telefonerna som såldes led 13 st av antennfelet. Är det statistisk säkerställt att andelen tillverkade iDroid-telefoner som lider av antennfelet är större än en procent? (6p)

4. Ett säljargument för nya iDroid är att batteritiden för telefonen är längre än den hos de finska konkurrenternas telefoner. De finska konkurrenterna tror inte på detta och mätte därför batteritiderna (i standbyläge) för 500 iDroid och 500 telefoner av deras eget märke. För iDroid fick de $\bar{x}_1 = 4560$ minuter och $s_1 = 780$. För de egna telefonerna fick de $\bar{x}_2 = 4390$ minuter och $s_2 = 822$.
   (a) Använd ett 95 % konfidensintervall för att avgöra om iDroid verkligen har längre batteritid. (5p)
   (b) De finska ingenjörerna misstänker utifrån tidigare erfarenhet att batteritiden för mobiltelefoner inte är normalfördelad. Kan de ändå använda konfidensintervallet som beräknades i (a)? Varför eller varför inte? (3p)

B-delen från tentan 2011-08-26

6. Triangelfördelningen används ibland för att modellera storheter vars största respektive minsta värde är känt och är vanlig inom statistiskt finans- och projektledningsarbete. En slumpvariabel X sägs vara triangelfördelad på intervallet (0,1) om dess täthetsfunktion är

$$f(x) = 2 - 4|x - 0.5|, \qquad 0 < x < 1.$$

   (a) Beräkna $E(X)$ och $V(X)$. *Ledning: att först skissa täthetsfunktionen kan kanske underlätta beräkningarna här.* (3p)
   (b) Beräkna sannolikheten att summan av 100 sådana oberoende slumpvariabler är större än 50. *Observera att du delvis kan lösa (b) även om du inte lyckas lösa (a).* (3p)

B-delen från tentan 2011-08-26

5. Ett fönster anses vara klimatsmart om dess U-värde är mindre än 1.2. U-värdet är ett mått på hur väl en del av en byggnad isolerar mot värmeförluster och mäts i W/m^{2}K. För ett fönster på 1 m^2 innebär ett U-värde på 1.2 alltså att fönstret läcker ut 1.2 Watt per Kelvin temperaturskillnad mellan ute och inne. En fönstertillverkare mätte värmeförlusten för ett 1 m^2 stort fönster vid olika temperaturskillnader:

Temperaturskillnad x (K)	1.00	3.00	4.00	5.00	8.00
Värmeförlust y (W)	0.30	2.42	5.74	6.20	9.39

Detta kan summeras med följande värden: $\bar{x} = 4.2$, $\bar{y} = 4.807$ $S_{xx} = 26.8$, $S_{yy} = 49.860$, $S_{xy} = 35.641$. Värmeförlusterna kan antas vara normalfördelade, oberoende och ha samma varians. Vi antar modellen $y = kx + m + \epsilon$

(a) Vilken parameter i modellen motsvarar U-värdet? (1p)

(b) Skatta fönstrets U-värde. (1p)

(c) Beräkna ett 95 % konfidensintervall för U-värdet. Kan fönstret anses vara klimatsmart? (3p)

B-delen från tentan 2011-08-26

3. I en fabrik som tillverkar komponenter till solceller tog man ett stickprov bland de tillverkade enheterna för att avgöra om de håller den kvalitet som man önskar. Av 150 undersökta enheter så fann man att 10 stycken var undermåliga. Företagsledningen kräver att andelen undermåliga enheter i produktionen ska vara mindre än 10 %.

   (a) Skatta andelen undermåliga enheter i produktionen. (1p)

   (b) Ange medelfelet för skattningen i (a). (2p)

   (c) Beräkna ett konfidensintervall för andelen undermåliga enheter och använd det för att svara på om företagsledningens krav är uppfyllt eller inte. (3p)

B-delen från tentan 2010-08-25

2. Låt slumpvariabel X beskriva den årliga maximala våghöjden till havs, enhet meter. Man har funnit att X följer en s.k. Gumbelfördelning med fördelningsfunktionen

$$F_X(x) = e^{-e^{-(x-b)/a}}, \qquad -\infty < x < \infty$$

där a och b är parametrar med värdena $a = 3$ och $b = 12$. Beräkna sannolikheten för en årlig maximal våghöjd högre än 14 meter. (6p)

3. I en studie undersöktes reparationstiderna för maskinfel av två slag. Från insamlade data har man funnit 58 fel av första slaget, vilka tog i genomsnitt 79.7 minuter att reparera (standardavvikelse 18.4 minuter). Vidare förekom 71 fel av andra slaget, med en genomsnittlig reparations om 87.3 minuter (standardavvikelse 19.5 minuter).
Beräkna ett 99 % konfidensintervall för skillnaden i genomsnittlig reparationstid mellan de två typerna av maskinfel. Kommentara det erhållna konfidensintervallet. (6p)