Matematisk statistik
Matematikcentrum
Lunds tekniska högskola, Lunds universitet

Tentamen: 2012–12–18 kl 8^{00}–13^{00}
FMS 012 — Matematisk statistik för PiE, F, CDI, 9 hp
MAS B03 — Matematisk statistik för fysiker, 9 hp

Korrekt, väl motiverad lösning på uppgift 1–3 ger 10 poäng vardera medan uppgift 4–6 ger 20 poäng vardera. Totalt kan man få 90 poäng. Gränsen för godkänd är 40 poäng.

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

Tillåtna hjälpmedel: Matematiska och statistiska tabeller som ej innehåller statistiska formler, Formelsamling i matematisk statistik AK 2001 eller senare, samt miniräknare.

Resultatet läggs in i LADOK *senast* 2012-12-28. Det anslås även på kurshemsidan.

---

1. (a) För de två s.v. X och Y gäller att $E(X) = 1$, $E(Y) = 2$, $V(X) = 5$, $V(Y) = 4$ samt $\rho(X, Y) = -0.5$. Beräkna väntevärde och varians för den s.v. $3X + 2Y$. (3p)

   (b) Mängden aktiv substans (mg) i en tablett beskrivs enligt en normalfördelning, $N(2, \sigma)$. Om mängden aktiv substans understiger 1.9 mg anses tabletten inte vara tillräckligt bra för medicinering. Hur stort får σ högst vara om högst 5% av tabletterna får ha denna defekt? (4p)

   (c) Den s.v. X har täthetsfunktion
   $$f_X(x) = (x + 1)^{-2}, \;\; x \geq 0.$$
   Bestäm sannolikhetsfunktionen för heltalsdelen av X, dvs för $Y = [X]$. (3p)

2. I uppsatsen ”Effect of Refrigeration on the Potassium Bitartrate Stability and Composition of Italian Wines” av A. Versari et al., *Italian Journal of Food Science* 2002:45-52) diskuteras mängden vinsyra i åtta olika viner som har utsatts för en kylstabiliseringsprocess. Nedan finns mätningar av mängden vinsyra (g/l) i de åtta vinerna före och efter kylbehandlingen.

Vin	1	2	3	4	5	6	7	8
Före	2.86	2.85	1.84	1.60	0.80	0.89	2.03	1.90
Efter	2.59	2.47	1.58	1.56	0.78	0.66	1.87	1.71

   Undersök, med ett lämpligt test på nivå 5%, om kylbehandlingen sänker vinsyran i vinet. Normalfördelning kan förutsättas. (10p)

3. Den tvådimensionella stokastiska variabeln (X, Y) har täthetsfunktionen
   $$2e^{-x}e^{-2y}, \quad x \geq 0, y \geq 0$$
   Bestäm $P(X > Y)$. (10p)

4. Antalet glassar som säljs i en liten kiosk en viss sommardag är Poissonfördelat med ett väntevärde m som beror på vädret. Soliga dagar är $m = 30$, mulna dagar är $m = 10$ och regniga dagar är $m = 2$. Vädret varierar enligt en Markovkedja med tillstånden E_1: ”soligt”, E_2: ”mulet” och E_3: ”regnigt”, med övergångsmatris
   $$P = \begin{pmatrix} 0.7 & 0.3 & 0 \\ 0.2 & 0.4 & 0.4 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

   (a) Antag att det är mulet idag. Beräkna sannolikheten att man inte får sålt en enda glass idag. (4p)

   (b) Antag att det är mulet idag. Beräkna sannolikheten att man inte får sålt en enda glass i morgon. (4p)

   (c) Beräkna den asymptotiska fördelningen för vädret. (6p)

   (d) Beräkna väntevärdet för antalet sålda glassar en dag mot slutet av sommaren. (6p)

5. Financial Times Stock Exchange (FTSE-100) är ett index som omfattar de 100 största aktierna, mätt till marknadsvärde, på Londonbörsen. Man vill nu undersöka om man kan förklara avkastningen för FTSE-100 (y) med förklaringsvariablerna

x_1	=	avkastningen på den brittiska obligationsmarknaden
x_2	=	avkastningen på S&P 500 (index inkluderar ett representativt urval av större ledande bolag på den amerikanska marknaden)
x_3	=	avkastningen på växlingskursen mellan amerikanska dollar och brittiska pund.

   Under 28 dagar mätte man värdet på y och x_1, x_2, x_3.

   (a) Data har först analyserats enligt följande modell:
   Modell 1: $Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ där $\varepsilon_i, i = 1, \ldots 28$, är oberoende och $N(0, \sigma_1)$.
   Från analysen sammanställdes följande tabell över skattade parametrar och dess medelfel

i	β_i^*	$d(\beta_i^*)$
0	-0.3420	0.6810
1	-0.0613	0.2644
2	0.9390	0.1504
3	-0.0328	0.1908

   Undersök, med lämpliga test eller intervall, om var och en av förklaringsvariablerna gör nytta (d.v.s. om den bör vara med i modellen). Varje test ska vara på signifikansnivå 5%. (5p)

   (b) Data har sedan analyserats enligt följande modell:
   Modell 2: $Y_i = \beta_0 + \beta_2 x_{i2} + \varepsilon_i'$ där $\varepsilon_i', i = 1, \ldots 28$, är oberoende och $N(0, \sigma_2)$.
   Från analysen: $Q_0 = 253.09$,

i	β_i^*	$d(\beta_i^*)$
0	-0.4258	0.5954
2	0.9359	0.1429

   $$(X^T X)^{-1} = \begin{pmatrix} 0.0364 & -0.0012 \\ -0.0012 & 0.0021 \end{pmatrix}$$

   Vilken avkastning kan vi förvänta oss för en genomsnittsdag för FTSE-100 när avkastningen för S&P 50 är 3%, d.v.s. $x_2 = 3$. Bilda ett 95% konfidensintervall för lämplig storhet. (10p)

   (c) Utgående från modell 2, punktskatta skillnaden i avkastning för FTSE-100 för två dagar då avkastningen för S&P 500 är 2% och 4%. (5p)

6. Weibullfördelningen är en av de mest använda fördelningarna för att beskriva livslängder av olika slags komponenter. Den stokastiska variablen X är Weibullfördelad om

   $$P(X > x) = e^{-\frac{1}{a} \cdot x^c}, x \geq 0$$

   där a och c är givna positiva konstanter.

   (a) Härled täthetsfunktionen till X. (2p)

   (b) För en viss sorts komponenter är $c = 2$, medan a är okänd. Man gör därför observationer av livslängderna på 5 komponenter och får värdena (år): 0.26 0.30 0.34 0.74 0.95.
   Beräkna maximum-likelihoodskattningen av a. (10p)

   (c) Skatta på lämpligt sätt percentilen L_{10} för komponenterna genom att utnyttja resultatet i (b). Med L_{10} menas det värde som uppfyller $P(X \leq L_{10}) = 10\%$. Om du inte lyckats få någon skattning av a i (b) kan du använda att $a^* = 0.345$. (3p)

   (d) Undersök om maximum-likelihoodskattningen från (b) är väntevärdesriktig. Om X är Weibullfördelad med $c = 2$ gäller att $E(X) = \sqrt{\frac{\pi}{4} \cdot a}$ och $V(X) = \frac{4-\pi}{4} \cdot a$. (5p)

Lycka till!
Glöm inte svara på CEQ-enkäten som mailats till dig!