

**TENTAMEN:** Matematisk statistik och diskret matematik (MVE050/MSG810). Statistik för fysiker (MSG820)

**Tid och plats:** Fredagen den 17 december 2010, kl. 8.30–12.30, Väg och vatten.

**Jour:** Stefan Erikshed, tel 0736281687.

**Tillåtna hjälpmedel:** Chalmersgodkänd räknare och Beta.

**Betygsgränser:** Chalmers: 3: 12 poäng, 4: 18 poäng, 5: 24 poäng. GU: G: 12 poäng, VG: 21 poäng. Maximalt antal poäng är 30.

1. (3p) Grundläggande sannolikhetslära.

- Definiera vad som menas med att två händelser är disjunkta respektive oberoende.
- Antag att andelen bilister som är alkoholpåverkade en lördagskväll är 0.4%. En viss typ av alkotest ger utslag i 99% av fallen då testpersonen verkligen är berusad. Sannolikheten är även 0.6% att testet ger utslag givet att testpersonen är nykter. Polisen stoppar en bilist en lördagskväll och alkotestet ger utslag, hur stor är sannolikheten att bilisten verkligen är berusad?

2. (3p) Anropen till en viss webserver kan antas följa en Poissonprocess med intensiteten 100 anrop i timmen.

- Vad är sannolikheten att det under en period på 1 minut sker mer än 1 anrop?
- Vad är sannolikheten att det efter ett anrop sker ännu ett inom 5 sekunder?

3. (3p) Grundämnet fluor är giftigt i för stora mängder. På en viss ort tas ett stickprov av fluorhalten i dricksvattnet, med följande mätvärden;

1.34 1.32 1.25 1.37 1.42 [mg/liter]

- Beräkna ett dubbelsidigt konfidensintervall för medelhalten,  $\mu$ , med konfidensgraden 95%. Du kan anta att mätvärdena är normalfördelade.
- Livsmedelsverket vill undersöka om medelhalten i orten överskrider 1.3 mg/liter, kan du avgöra om så är fallet?

4. (4p) I en slantsingling väljer Eva klave och förlorar. Eva misstänker att myntet är orättvist. Hon ställer upp följande hypotestest för sannolikheten att få klave,  $p$ .

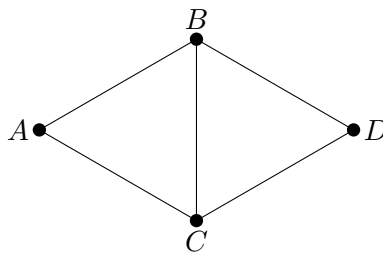
$$H_0 : p = 0.50$$

$$H_1 : p < 0.50$$

Eva bestämmer dig för att singla myntet 20 gånger, och om hon får 6 eller färre klave så förkastar hon nollhypotesen.

- Vad innebär Typ I-fel och vad är risken för att Eva gör ett sådant i detta fall?
- Vad kallas risken för ett Typ I fel med ett annat ord och hur brukar man beteckna den?
- I verkligheten är myntet orättvist, sannolikheten att få klave är bara  $p = 0.30$ . Beräkna med hjälp av denna information risken att Eva gör ett Typ II-fel.
- Vad blir testets styrka?

5. (3p) Myran Ante kryper längs kanterna på grafen nedan, de fem kanterna är vardera 1 dm långa. När Ante kommer till ett hörn fortsätter han längs någon av kanterna som ansluter till hörnet med lika sannolikhet. Han väljer kant oberoende av hur han krupit tidigare. Han kan alltså även gå tillbaka längs samma kant som han nyss kom på. Hörn  $D$  är dock försett med klister, så att Ante fastnar när han kommer dit! Om Ante startar i hörn  $A$ , vad är väntevärdet för sträckan han kryper innan han fastnar i klistret?



6. (3p) Längs en husfasad på Magasinsgatan i Göteborg finns talföljden 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, ... nedtecknad. Den kallas för Fibonaccis talföljd och definieras genom rekursionsrelationerna

$$\begin{cases} F_k = F_{k-1} + F_{k-2}, & k \geq 2 \\ F_0 = 0 \text{ och } F_1 = 1. \end{cases}$$

Bestäm den genererande funktionen till Fibonaccis talföljd  $(F_k)_{k=0}^{\infty}$ .

7. (4p) Mats konstruerar ett nytt spamfilter till en mailtjänst. Vid ett test så visar det sig att filtret stoppar 931 av 950 spam-mail.
- Gör en punktskattning av filtrets tillförlitlighet,  $p$ , d.v.s. andelen spam som filtret lyckas stoppa.
  - Beräkna ett enkelsidigt konfidensintervall för  $p$  av grad 95%, d.v.s. ett konfidensintervall för  $p$  på formen  $[L, 1]$ .
  - Vilka approximationer används vid härledningen av ett sådant konfidensintervall?
  - Mailtjänstens gamla spamfilter hade en tillförlitlighet på 97.15%. Kan Mats hävda att hans nya filter är bättre?
8. (3p) Ida planerar att köpa en begagnad bil och vill uppskatta vilken reparations- och underhållskostnad hon kan räkna med under det första året. Från en bilhandlare hon känner får hon veta årskostnaden för fem olika bilar tillsammans med deras körsträcka.

Körsträcka [1000 mil]	13.0	5.0	8.5	7.2	19.0
Kostnad [kr]	6700	900	3600	2200	14000

- Skatta en enkel linjär regressionsekvation till dessa data, där du låter körsträckan  $X$  vara den förklarande (deterministiska) variabeln och kostnaden  $Y$  responsvariabeln (stokastisk).
- Ida väljer mellan två bilar i olika prisklasser, den ena har gått 15000 mil och den andra 38000 mil. Kan du använda din modell för att skatta hennes första årskostnad i de två fallen?

9. (4p) Stora talens lag.

- a) Formulera Chebyshevs olikhet, med premisser och förklaring av beteckningar.
- b) Bevisa följande sats.

*Sats:* Låt  $X_1, X_2, \dots, X_n$  vara oberoende och likafördelade stokastiska variabler med  $\mathbf{E}(X_i) = \mu$ , där  $|\mu| < \infty$ , och  $\mathbf{Var}(X_i) = \sigma^2 < \infty$ . Låt vidare  $S_n = \sum_{i=1}^n X_i$ . Då gäller  $\forall \epsilon > 0$

$$P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) \rightarrow 1 \text{ då } n \rightarrow \infty.$$

Lycka till och God Jul!