TENTAMEN: Matematisk statistik och diskret matematik D (MVE055/MSG810)
Tid och plats: Tisdagen den 24 augusti 2010, kl. 08.30–12.30, VV-salar.
Jour: Oscar Hammar, tel. 0708-300715.
Tillåtna hjälpmedel: Chalmersgodkänd räknare och Beta.
Betygsgränser: 3: 12 poäng, 4: 18 poäng, 5: 24 poäng. Maximalt antal poäng är 30.

1. (3p) Låt X och Y vara två oberoende normalfördelade stokastiska variabler, där X har väntevärde 2 och standardavvikelse 1, medan Y har väntevärde -7 och standardavvikelse 3.

   a) Bestäm fördelningen (inklusive parametrar) för $2X + Y$.
   b) Vad är $\mathbf{P}(5X > 10)$?

2. (3p) Bestäm $\mathbf{P}(X \geq 3)$ om X är

   a) antalet kast med en (vanlig ideal sexsidig) tärning tills dess att en tvåa kommer upp.
   b) antalet ja-svar i statistisk undersökning där ett slumpmässigt stickprov om 20 personer tillfrågas i en viss fråga, om andelen som skulle svarat ja i hela populationen är 20%.
   c) summan av kvadraterna på två oberoende standardnormalfördelade variabler, dvs $X = Y^2 + Z^2$ där $Y, Z \sim \text{Norm}(0, 1)$.

3. (3p) Erik har noterat att 40% av hans epostmeddelanden är skräppost. Av dessa innehåller hela 25% ordet "lottery". Detta ord verkar vara en väldigt bra indikator på vad som är skräppost, eftersom endast 15% av samtliga hans epostmeddelanden innehåller just det ordet.

   a) Vad är sannolikheten för att ett slumpmässigt epostmeddelande som innehåller ordet "lottery" är skräppost?
   b) Hur stor andel av de epostmeddelanden som ej är skräppost innehåller ordet "lottery"?

4. (3p) I syfte att avgöra om det är en fördel eller en nackdel att spela fotboll på hemmaplan valdes 150 icke oavgjorda matcher ut slumpmässigt. Bland dessa visade sig 83 matcher ha vunnits av hemmalaget.

   a) Bestäm ett 99%-igt dubbelsidigt konfidensintervall för andelen icke oavgjorda matcher som vinns av hemmalaget.
   b) Vilken slutsats kan du baserat på konfidensintervallet dra om påståendet *"det är större sannolikhet att vinna än förlora om man spelar på hemmaplan"*.

5. (4p) Lastbilschaufförern Beata har fått en fix idé: hon samlar på nummer på registreringsskyltar - *i stigande ordning*. Det hela började med hon såg en bil med registreringsnummer "GAH 001" och strax därefter en med "HOI 002". Sedan letade hon länge efter en vars sifferdel var "003" och så vidare.

   a) Hennes mål är att bocka av alla nummer från "001" till "999" i stigande ordning. Vad är väntevärdet av antalet registreringsskyltar hon måste se innan hon är klar?
   b) Letandet kan ibland ta tröstlöst lång tid, men en gång hade hon tur; just då hon mött en bil med det nummer hon letade efter såg hon att bilen bakom denna hade det efterföljande numret och kunde därför bocka av två nummer på en gång. Detta verkade väldigt turligt, men var det verkligen det? Vad är sannolikheten att detta skulle inträffa någon gång under letande efter skyltar med sifferdelar från "001" till "999"?

   Obs: Du kan utgå från att hon bara läser av en skylt åt gången. Dessutom kan du som en god approximation anta att alla nummer mellan "001" och "999" är lika vanliga och att sifferdelen på en skylt inte på något sätt påverkas av sifferdelen på tidigare avlästa skyltar.

6. (4p) I ett digital optiskt kommunikationssystem där sändaren består av en laserdiod och mottagaren av en fotonräknare kan antalet räknade fotoner under sändningen av en enskild bit antas vara Poissonfördelad med väntevärdet 3.0 om den sända biten var en 0:a, medan väntevärdet är 10.0 om den sända biten var en 1:a. I mottagaren tolkas biten som en 0:a om antalet räknade fotoner var 6 eller färre. Annars tolkas den som en 1:a.

   a) Om en 0:a sänds, vad är sannolikheten att den tas emot felaktigt (dvs tolkas som en 1:a av mottagaren)?
   b) Anta att 0:or och 1:or sänds med lika stor sannolikhet. Om antalet räknade fotoner var exakt 6, vad är då sannolikheten att en 0:a sändes?

7. (4p) Antalet sökningar under en slumpmässigt vald minut hos en större sökmotor på internet kan med god approximation antas vara normalfördelat. Under 9 olika slumpmässigt utvalda minuter registrerades följande antal sökningar:

$7.13 \cdot 10^5$	$6.82 \cdot 10^5$	$6.10 \cdot 10^5$	$5.42 \cdot 10^5$	$9.16 \cdot 10^5$
$4.29 \cdot 10^5$	$5.99 \cdot 10^5$	$6.28 \cdot 10^5$	$5.13 \cdot 10^5$	

   Bestäm 95%-iga konfidensintervall för väntevärdet och standardavvikelsen för antal sökningar under en slumpmässigt utvald minut.

8. (3p) Låt X vara en diskret stokastisk variabel som antar värdet $2k$ med sannolikheten $\frac{1}{2^k}$ för alla positiva heltal $k = 1, 2, 3, \ldots$.
   Beräkna den momentgenererande funktionen $m_X(t)$ för X och ange för vilka värden på t som den existerar.

9. (3p) Den kontinuerliga stokastiska variabeln Y har täthetsfunktionen

$$f_Y(y) = \frac{2y}{p}, \quad 0 \leq y \leq \sqrt{p}$$

   där $p > 0$ är en okänd parameter. Låt $Y_1, \ldots, Y_n$ vara ett slumpmässigt stickprov av Y och bestäm konstanten a så att $a\frac{1}{n}\sum_{i=1}^{n} Y_i^2$ är en väntevärdesriktig skattare av p.

Lycka till!