**UPPSALA UNIVERSITET**
**Matematiska institutionen**
Fredrik Jonsson
jonsson@math.uu.se

MATEMATISK STATISTIK
Tentamen för NVP kemi/geologi
Matematik och statistik NV1
Statistikdelen
2007-01-11

*Skrivtid: 9-13. För godkänt krävs 18 poäng, för väl godkänt 28 poäng (inklusive bonuspoäng). Tillåtna hjälpmedel: räknedosa och utdelad Formel- och Tabellsamling. Lösningar skall vara väl motiverade och försedda med förklarande text. Om fullständig lösning saknas, försök ange, i ord och med resultat, hur långt du har kommit och vad du har tänkt. Lycka till!*

1. Oskar och Jonas tycker om att spela sällskapsspel. I ett av spelen ingår, förutom den vanliga 6-sidiga tärningen, också 8-sidiga och 20-sidiga tärningar.

   (3p) Vid ett tillfälle ska Oskar slå tre stycken 8-sidiga tärningar (så kallade "T8-or"). För att lyckas hoppas han att åtminstone en av tärningarna visar en åtta. Han tror att chansen att lyckas är 3/8, dvs. knappt 40%. Kan du bekräfta eller annars räkna ut den korrekta sannolikheten åt honom?

   Vid ett annat tillfälle slår Jonas tre stycken vanliga "T6-or" och lägger ihop resultaten.

   (2p) Vad är sannolikheten att få resultatet 17 eller högre, dvs. 17 eller 18?

   (2p) Om man jämför denna procedur med att slå en T20, vilket ger då i genomsnitt högst värde? Ange ditt svar genom att beräkna relevanta väntevärden.

2. Ett piratföretag tillverkar glödlampor av undermålig kvalitet: 15% av alla lampor är defekta. För att undvika alltför omfattande klagomål kompletterar man varje "10-förpackning" med två bonuslampor, så att de i själva verket innehåller 12 lampor.

   (2p) Vad är sannolikheten att en förpackning med 12 lampor innehåller åtminstone 10 felfria exemplar?

   (2p) Åsa köper 20 förpackningar med 12 lampor i varje. Vad är sannolikheten att totalt högst 30 lampor är defekta?

   (2p) Vid ett mera välkänt företag är 99.5% av lamporna felfria. Annika köper 24 förpackningar med 10 lampor i varje (inga bonusexemplar här). Vad är sannolikheten att fler än två av de inköpta lamporna är defekta?

3. Man tänker sig ett sannolikhetsrum som rymmer möjligheterna av ett obegränsat antal slantsinglingar (krona eller klave i varje försök). Det fungerar alltså som modell för situationen då man genomför upprepade, oberoende experiment som antingen lyckas eller misslyckas med samma sannolikhet varje gång. Vi tänker oss dessutom att chansen att lyckas inte behöver vara 1/2 utan mer generellt $p$.

   Man kan då introducera en slumpvariabel $X$ som anger antalet misslyckade försök före det första lyckade. Rent teoretiskt är det förstås möjligt att alla försök misslyckas "ända in i oändligheten", men man kan visa att sannolikheten att detta ska ske är 0. Därför kan man anse att $X$ är en väldefinerad slumpvariabel och inom sannolikhetsteorin säger man att $X$ är *geometriskt fördelad med parameter* $p$.

   (1p) Vilka värden kan $X$ anta? Vad är sannolikheten att $X = 0$?

   (3p) Ange sannolikheten att $X = k$ som en funktion av $p$ och $k$. Börja med små värden på $k$ och försök sedan hitta ett allmänt uttryck.

(3p) Studenten Isac tycker att en viss kurs har varit tråkig och har därför inte lagt ner så mycket jobb före tentan. Han säger att det är "fifty-fifty" om han ska klara den eller ej. Om han misslyckas så kan han tänka sig att försöka igen vid ytterligare två tillfällen. Han tror inte att hans chanser påverkas på något sätt efter eventuella misslyckanden. Vad är sannolikheten att Isac får med sig kurspoängen?

4. Under en studie av kopparormar samlar man in fem fullvuxna han- respektive hondjur och mäter längden av dessa. Resultat (i cm):

| | | | | | |
|---|---|---|---|---|---|
| *Hanar* | 22.07 | 23.03 | 21.57 | 22.28 | 22.37 |
| *Honor* | 21.57 | 21.89 | 22.24 | 22.52 | 22.02 |

De båda längdserierna kan antas utgöra oberoende stickprov från normala fördelningar med samma varians. Man vill bl.a. veta om det är någon skillnad mellan hanars och honors genomsnittslängd.

(3p) Bilda ett 95% konfidensintervall för differensen mellan väntevärdena.

(1p) Tyder resultatet i första frågan på att det föreligger en skillnad mellan hanars och honors genomsnittslängder? Motivera.

(3p) En naturboksförfattare skulle vilja ha ett hälften så brett intervall. Hur många observationer skulle detta kräva? Vi förutsätter att lika många hanar som honor väljs ut, samt att stickprovsvarianserna förblir oförändrade.

5. Vid en sociologisk undersökning (i USA) fick 1397 personer redogöra för sin inställning till vapenlicens och dödsstraff. Resultatet kan sammanfattas så här:

| | *För vapenlicens* | *Mot vapenlicens* | *Summa* |
|---|---|---|---|
| *För dödsstraff* | 784 | 311 | 1095 |
| *Mot dödsstraff* | 236 | 66 | 302 |
| *Summa* | 1020 | 377 | 1397 |

(7p) Testa (med hjälp av ett lämpligt $\chi^2$-test) på 5%-signifikansnivå om inställningen till vapenlicens kan anses oberoende av inställningen till dödsstraff. Vilka slutsatser kan man dra?

6. Två mattelärare, A och B, gillar båda att spela boule. Under en sommar, då tillvaron är lite lugnare och området kring Polacksbacken och Ångströmlaboratoriet är så gott som studentfritt, händer det rätt ofta att de avslutar dagen med att gå ut i den ljumma kvällen för att spela några omgångar. De för förstås lite statistik över hur det går, nedan visas resultatet (50 omgångar).

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | -2 | -2 | -1 | 5 | -7 | 10 | -1 | -1 | -1 | -3 | -2 | 3 | 4 | -1 | 5 | -2 |
| -2 | 5 | -1 | -1 | 4 | 5 | -3 | -2 | 12 | 3 | -4 | 4 | 4 | -3 | 11 | -1 | 7 |
| -1 | -1 | -2 | 4 | 10 | -1 | 2 | -3 | 8 | 7 | -1 | -2 | 6 | -2 | 1 | 1 | |

I varje omgång gäller det att komma först till 13 och efter omgångens slut antecknar de poängskillnaden. Negativa värden innebär att A förlorat, positiva att A vunnit. När sommaren börjar gå mot sitt slut är de oense om vem som egentligen är den bättre spelaren. Kan du hjälpa till med en enkel dataanalys?

(4p) Redovisa datamaterialet grafiskt och kommentera. Det kan vara lämpligt att göra ett stolpdiagram och/eller ett lådagram. Hur skulle du beskriva spelarna och resultatet utifrån detta? Vem utav dem är den bättre?

(2p) Beskriv kortfattat vad som menas med ett stolpdiagram respektive ett lådagram. Vad finns det för skillnader i den information de ger? Finns det för- och nackdelar?