

Tentamen, Linköpings universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn: 732G70 Statistik A och HSTA75 Grundläggande statistik och matematik
Datum och tid: 2008-04-26 kl 08-12
Jourhavande lärare: Olle Eriksson 1437
Tillåtna hjälpmedel: Miniräknare, tabeller (9 sidor för 732G70, 8 sidor för HSTA75),
Statistisk dataanalys av Körner/Wahlgren, 3:e eller 4:e upplagan.
Boken får innehålla anteckningar.

Uppgifterna är helt påhittade och sifferuppgifterna är inte säkert i närheten av verkliga värden.

Uppgift 1—4 är påhittade, men bakgrunderna är delvis hämtade från presentationer av aktuella undersökningar i dagstidningar, nyhetssändningar och på texttv 2008-04-23.

1. (5 poäng) I en undersökning har man funnit ett ämne i tröjor hos några av de vanliga klädkedjorna, ett ämne som kan vara giftigt vid stora doser. Antag att man tagit ett stickprov om 10 röda tröjor från ett mycket stort parti och registrerat följande mängder (gram) av detta ämne. Det insamlade materialet visas här:

0.53 0.53 0.32 0.35 0.28 0.29 0.37 0.23 0.14 0.24

- (a) Beräkna stickprovets medelvärde och standardavvikelse.
(b) Skatta μ med 95% konfidensintervall.
(c) Förklara tydligt hur konfidensintervallet ska tolkas.
(d) Förklara tydligt vilka antaganden du måste göra för att ditt svar ovan ska vara giltigt.
2. (3 poäng) Se bakgrund från uppgift 1. Antag nu att man tagit ett motsvarande stickprov av blåa tröjor och ett stickprov av svarta tröjor. Data sammanställs i följande tabell:

färg	n	\bar{x}	s
blå	12	0.60	0.10
svart	17	0.50	0.12

- (a) Testa på 5% risknivå $H_0 : \mu_b = \mu_s$, $H_1 : \mu_b > \mu_s$
(b) Blå färg innehåller lite mer av det aktuella ämnet. Man har haft uppfattningen att blåa tröjor i genomsnitt innehåller 0.02 gram mer än svarta. Testa på 5% risknivå nollhypotesen att den gamla uppfattningen är korrekt mot alternativhypotesen att den gamla uppfattningen är felaktig.
3. (5 poäng) Organisationen RoF (Restauranger och Frisörer) har 1870 småföretag som medlemmar, 1120 restauranger och 750 frisörer. För att få en uppfattning om läget bland de egna medlemmarna ställer man frågan om de fått betala straffavgift efter Skatteverkets kontroll av personalliggare. Av 120 slumpmässigt utvalda restauranger svarar 15 ja och av 100 frisörer svarar 10 ja.
- (a) Skatta *antalet* som fått betala sådan straffavgift bland alla anslutna *restauranger* med 95% konfidensintervall.
(b) Skatta *andelen* som fått betala sådan straffavgift bland alla anslutna *företag* med 95% konfidensintervall.

4. (2 poäng) Man har gjort en undersökning av mängden vatten i kycklingfileer. Antag att man kontrollerat 200 förpackningar och kommit fram till att det är för mycket vatten i 26 av dem, vid jämförelse mot en gräns man tidigare har kommit överens om.
- En person påstår att ett konfidensintervall för π är 0.10—0.16. Vilken konfidensgrad har man i detta yttrande?
 - Man betraktar undersökningen enligt ovan som en provundersökning för att få ett bra underlag för att dimensionera en ny studie. Målet med den nya studien är att skatta π med 95% konfidensgrad med en felmarginal på 0.02. Hur stort stickprov bör man planera att dra i den nya studien?
5. (1 poäng) På ett nöjesfält finns ett lyckohjul med 20 fält. Vinst delas ut om hjulet stannar på det fält man har satsat på. Varje fält antas ha lika stor sannolikhet och det råder oberoende mellan spelomgångarna. I varje omgång ska du satsa på ett fält. Beräkna sannolikheten att du ska vinna exakt 2 gånger om du är med under 10 spelomgångar.
6. (4 poäng) Antalet skräpbrev, X , som kommer till en viss användare varierar slumpmässigt från dag till dag på ett sätt som beskrivs av följande sannolikhetsfördelning:

x	$P(X = x)$
0	0.20
1	0.40
2	0.30
3	0.10

- Beräkna fördelningens väntevärde och varians.
- Skräppost rensas automatiskt av ett speciellt program, som vi antar fungerar precis som det är tänkt d.v.s. det rensar all skräppost men släpper igenom all riktig post. Enligt licensvillkoren betalar man 50 öre per dag i fast avgift och därutöver 8 öre per skräpbrev. Brev som inte sorteras som skräp kostar inget extra, och det motiveras av att de inte ger någon belastning i form av efterkontroller, uppdatering av spärrlistor m.m. Beräkna väntevärde och varians för kostnaden för programlicensen under en dag.

Inför de två följande deluppgifterna gör vi antagandet att det inte finns något beroende mellan dagar. Räkna också med att ett år har 365 dagar.

- Beräkna väntevärde och varians för antal skräpbrev som kommer under ett år.
- Beräkna sannolikheten att det kommer mer än 500 skräpbrev under ett år.