

TENTAMEN I STATISTIKENS GRUNDER 2 2011-03-18

Skrivtid: 9.00-14.00

Godkända hjälpmedel: Miniräknare, språklexikon

Tentamen består av fem uppgifter. För full poäng på en uppgift krävs tydliga, utförliga och väl motiverade lösningar.

OBS! Glöm inte att ange nödvändiga antaganden överallt.

Uppgift 1. (20 poäng)

Det ställs höga krav på säkerhet och kvalitet vid läkemedelstillverkning. Som ett led i kvalitetssäkringen görs regelbundet kvalitetskontroller av tillverkningsprocessen. Vid en sådan kontroll har 250 slumpmässigt valda tabletter analyserats. Innehållet av den verksamma substansen i preparatet ska vara 500 mg. Det genomsnittliga innehållet bland tabletterna i urvalet blev 499.9 mg. Antag att populationsvariansen är 1.2.

- a) Testa om det genomsnittliga innehållet avviker från 500 mg på signifikansnivån 1 %.
- b) Beräkna och tolka p-värdet för testet i a).

Uppgift 2. (20 poäng)

I undersökningen "Livsstil och kulturvanor" har ett slumpmässigt urval om 400 vuxna (minst 20 år gamla) invånare i Stockholm dragits. Respondenterna fick svara på frågor om i vilken utsträckning de har ägnat sig åt olika fritidsaktiviteter, bland annat hur ofta de har gått på teater. I tabellen nedan visas resultatet uppdelat på olika stadsdelar. Testa med hjälp av dessa data om det finns någon skillnad mellan olika stadsdelar när det gäller invånarnas teatervanor.

Antal teaterbesök/kvartal	Stadsdel			
	Södermalm	Kungsholmen	Norrmalm	Östermalm
Inga	54	45	50	40
1-2	36	32	28	29
>2	27	18	16	25

Uppgift 3. (20 poäng)

I undersökningen "Livsstil och kulturvanor" har ett slumpmässigt urval om 400 vuxna (minst 20 år gamla) invånare i Stockholm dragits. På frågan om, och i så fall hur ofta, man brukar träna på ett gym, svarade 68 stycken att de tränar på gym minst en gång per månad.

- a) Uppskatta proportionen vuxna invånare i Stockholm som tränar på gym minst en gång i månaden.
- b) Uppskatta variansen hos urvalsproportionen i uppgift a). Bör ändlighetskorrektion användas? Motivera!
- c) Gör ett 95 %-igt konfidensintervall för proportionen vuxna invånare i Stockholm som tränar på gym minst en gång i månaden.
- d) Ge en tolkning av konfidensintervallet i uppgift c).
- e) Hur stort urval bör man dra om man önskar att felmarginalen (halva konfidensintervallets bredd) i ett 95 %-igt konfidensintervall inte skall överstiga tre procentenheter? Man kan anta att proportionen vuxna i Stockholm som tränar på gym minst en gång i månaden inte överstiger 25 %.

Uppgift 4. (20 poäng)

Ett mindre skogsområde skall säljas genom anbud, dvs. de som vill köpa området lägger ett bud som hålls hemligt fram till en viss tidpunkt då de inlämnade buden offentliggörs. Den som har det högsta budet vid offentliggörandet vinner budgivningen och får köpa området för det pris som anges i budet. Gröna dalen AB bedömer att området kan ge en avkastning på 100.000 kr. VDn överväger att antingen ge budet 40.000 kr, budet 60.000 kr eller budet 80.000 kr. VDn bedömer också att konkurrenten Östers skogar kan ge ett bud på 50.000 kr, ett bud på 70.000 kr eller inte ge något bud alls. Ingen annan väntas ge något bud.

- a) Hjälp VDn för Gröna dalen att sätta upp beslutsmatrisen.
- b) Vilket bud bör VDn välja om maximinkriteriet används?
- c) Vilket bud bör VDn välja om sannolikheten att Östers skogar lägger budet 50.000 kr är 0.50 och sannolikheten att de lägger budet 70.000 kr är 0.25?

Uppgift 5. (20 poäng)

För att undersöka om två fastighetsmäklare, TB och SW, värderar lägenheter lika, lät man dem var för sig värdera nio lägenheter. Ange nödvändiga förutsättningar för att pröva hypotesen att TB och SW i genomsnitt värderar lägenheter lika. Använd signifikansnivån 5 % och genomför testet med de data som anges i tabellen nedan.

Lägenhet	Värdering TB	Värdering SW
1	1425000	1380000
2	2150000	2140000
3	2575000	2560000
4	5200000	5180000
5	1475000	1385000
6	5480000	5390000
7	7900000	7885000
8	3100000	3180000
9	2995000	2985000

Formler

Räknerregler för väntevärden och varianser där X och Y är stokastiska variabler, a , b och c är konstanter:

$$\begin{aligned}E(c) &= c \\E(cX) &= cE(X) \\E(c + X) &= c + E(X) \\E(aX + bY) &= aE(X) + bE(Y) \\V(c) &= 0 \\V(cX) &= c^2V(X) \\V(c + X) &= V(X) \\V(aX + bY) &= a^2V(X) + b^2V(Y) + 2abCov(X, Y)\end{aligned}$$

Väntevärde och varians för urvalsmedelvärdet $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ där alla $X_1, X_2, \dots, X_n$ är oberoende och har väntevärde μ och varians σ^2 :

$$\begin{aligned}E(\bar{X}) &= \mu \\V(\bar{X}) &= \frac{\sigma^2}{n}\end{aligned}$$

Ändlighetskorrektur:

$$\frac{N - n}{N - 1}$$

Urvalsvarians:

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

Testvariabler:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

$$Z = \frac{\bar{X} - \bar{Y} - D_0}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

$$t = \frac{\bar{X} - \bar{Y} - D_0}{\sqrt{S_p^2 \left(\frac{1}{n} + \frac{1}{m}\right)}}, \quad \text{där } S_p^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$$

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$