

Att tänka på (obligatorisk läsning)

- A. Redovisa Dina lösningar i en form som gör det lätt att följa Din tankegång. (Rättaren förutsätter att det dunkelt skrivna är dunkelt tänkt.). Motivera alla väsentliga steg i lösningen. Ange alla antaganden Du gör och alla förutsättningar Du utnyttjar
- B. Vid konfidensintervall måste Du dessutom ange vad intervallet avser att täcka samt teckna intervallet i symbolform innan de numeriska uppgifterna sätts in. Verbal slutsats av det framräknade resultatet krävs för full poäng.
- C. Vid signifikansanalys måste Du utöver vad som sagts ovan ange H_0 , H_1 , signifikansnivå (alt. P-värde), testfunktion (inkl. ev. antal frihetsgrader), resultat och verbal slutsats.
- D. Samtliga beräknade mått skall följas av en verbal slutsats för full poäng.
- E. Uppgifterna är inte ordnade efter svårighetsgrad.

UPPGIFT 1

Förklara de principiella skillnaderna mellan slumpmässigt och icke slumpmässigt urval. **(3p)**

UPPGIFT 2

Ett företag i telemarketingbranschen har under en tid studerat sambandet mellan antal timmars introduktionsutbildning (x) och prestation (y) bland nyanställda. Man har sedermera skatta en regressionskvation som ser ut på följande vis:

- A. Tolka innebörden av koefficienterna i ekvationen. **(3p)**
- B. Hur lång bör utbildningstiden vara om man vill uppnå en genomsnittlig prestation på 80? **(3p)**

UPPGIFT 3

I en undersökning tillfrågades 327 slumpmässigt utvalda personer om huruvida de plägar använda långkalsonger under vintern. 142 personer svarade ja på frågan.

- A. Beräkna ett 95 %-igt konfidensintervall för andelen långkalsongbärare i populationen. **(5p)**
- B. Man ansåg att konfidensintervallet var onödigt långt. Hur stort stickprov skulle behövas om man i en förnyad undersökning ville halvera konfidensintervallets längd? **(3p)**

UPPGIFT 4

Det sägs att sk GI-kost sänker blodsockerhalten generellt. I en studie fick 8 diabetiker under en försöksperiod följa en GI-optimerad meny. Deras genomsnittliga blodsockervärde mättes för de två veckorna närmast före försöksperioden och för de två veckor som försöket pågick:

Person nr	Genomsnittligt värde före försöket	Genomsnittligt värde under försöket
1	12	9
2	14	11
3	8	9
4	11	10
5	10	10
6	15	17
7	13	10
8	14	11

- A. Testa om resultatet i experimentet styrker hypotesen att GI-mat sänker blodsockerhalten generellt genom att genomföra ett t-test för parvisa (matchade) observationer på 5%-nivån. **(7p)**
- B. Antag att den genomsnittliga minskningen generellt är 2,5 och beräkna testets styrka. **(4p)**
- C. Ge förslag på ett icke-parametriskt test som man hade kunnat använda som alternativ till t-test. **(2p)**

UPPGIFT 5

Mina arbetar som kvalitetsansvarig på ett stort verkstadsföretag. Hon har kommit fram till att 25 % av alla personalrelaterade misstag som påverkar produktionen inträffar på måndagar och att 20 % inträffar under arbetsskiftets sista timme. Dessutom har denne kommit fram till att 4 % av de personalrelaterade misstagen uppträder under måndagens sista timme.

- A. Hur stor är sannolikheten att ett personalrelaterat problem som inträffar på en måndag inte inträffar den sista timmen på arbetsdagen? **(6p)**
- B. Är de båda händelserna ”misstag som inträffar på måndagar” och ”misstag som inträffar under den sista timmen” att betrakta som statistisk oberoende? **(2p)**
- C. Om det i genomsnitt inträffar 100 misstag i veckan, hur stor är sannolikheten att det under måndagens sista arbetstimme inträffar högst ett personalrelaterat misstag? **(6p)**
- D. Beräkna förväntad kostnad och standardavvikelse för kostnaden under ett år om man räknar med att verksamheten är i gång 46 veckor per år och att varje misstag i genomsnitt kostar företaget 800 kronor. **(4p)**

UPPGIFT 6

En studie som genomförts av försäkringskassan har tittat på i vilken utsträckning långtidssjukskrivna med diagnoser rörande nacke/rygg har lyckats med sin rehabilitering. Man har utgått från sitt register och studerat 1500 slumpmässigt utvalda personer med de aktuella diagnoserna. De har också haft minst en sammanhängande sjukskrivningsperiod som överstigit 90 dagar mellan åren 1998 och 2003.

Man har samlat in bland annat uppgifter om följande egenskaper:

- Ålder (år)
 - Kön (man/kvinna)
 - Ersättningsnivå före rehabilitering (25 %, 50 %, 75 % och 100 %)
 - Ersättningsnivå efter rehabilitering (25 %, 50 %, 75 % och 100 %)
 - Tid från första sjukskrivningsdag till initierat rehabiliteringsärende (dagar)
 - Tid från initierat rehabiliteringsärende till påbörjad rehabilitering (dagar)
 - Rehabiliteringens längd (dagar)
- A. Ange för varje variabel om den är kvalitativ eller kvantitativ och om den är diskret eller kontinuerlig. Ange också variabelns skalnivå. **(2p)**
- B. Ge förslag på hur man lämpligast kan använda deskriptiv statistik för att åskådliggöra var och en av dessa variabler univariat. **(4p)**
- C. Hur kan man använda deskriptiva metoder för att påvisa eventuella skillnader mellan män och kvinnor när det gäller handläggningen av rehabiliteringsärendet? **(3p)**
- D. Hur kan man åskådliggöra ett samband mellan ålder och utfallet av rehabiliteringen med hjälp av beskrivande statistik? **(3p)**