Tentamen, Linköpings universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn:	732G01, 732G02-A, 732G90-A, HSTA13, HSTA06-A, HSTA91-A, Statistik I del A
Datum och tid:	2008-06-13 klockan 08-12
Jourhavande lärare:	Olle Eriksson, 1437
Tillåtna hjälpmedel:	Statistisk dataanalys av Körner och Wahlgren 3:e eller 4:e upplagan *boken får innehålla anteckningar*, författarnas korrigeringar av fel i boken, miniräknare, tabeller (8 sidor).

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

1. (5 poäng) Tabellen nedan beskriver möjliga utfall och motsvarande sannolikheter för slumpvariabeln X :

x	$P(X = x)$
1.0	0.01
1.5	0.06
2.0	0.20
2.5	0.34
3.0	0.29
3.5	0.10

   (a) Beräkna väntevärde och varians hos X .

   (b) Y är medelvärdet av 100 oberoende observationer på X . Beräkna väntevärde och varians hos variabeln Y .

   (c) Beräkna $P(Y \leq 2.6)$, där Y definierades i föregående deluppgift.

   (d) Det finns en normalfördelad variabel, R, med okänt väntevärde men där det är känt att standardavvikelsen är exakt 8. Man vet också att $P(R < 41) = 0.27$. Använd detta underlag för att beräkna väntevärdet μ_R .

2. (5 poäng) I utredningen efter en tågolycka gör man prov på bärigheten hos ett slumpmässigt urval om 12 slipers från den aktuella sträckan. Följande värden registreras:

   35 42 45 47 48 52 50 53 45 46 49 49

   Vi behöver inte gå in på exakt vilket mått som används, utan kan nöja oss med att veta att högre bärighet representeras med ett högre värde.

   (a) Beräkna stickprovets medelvärde och standardavvikelse.

   (b) Man ska göra några olika kontroller. En kontroll avser att testa om den genomsnittliga bärigheten på den aktuella sträckan kan vara så hög som 50. Testa därför $H_0 : \mu = 50$, $H_1 : \mu < 50$ på 5% risknivå.

   (c) Man planerar en mer omfattande undersökning där man önskar att skatta μ med 95% dubbelsidigt konfidensintervall. Man siktar på att felmarginalen ska bli av storlek 0.8. Hur stort stickprov bör man då ta? Använd uppgifterna ovan som om det var en provundersökning för att dimensionera den mer omfattande undersökningen.

3. (3 poäng) Se bakgrund i tidigare uppgift. På en viss del i järnvägsnätet finns slipers av två olika typer. Man drar ett stickprov av slipers av den ena typen och ett annat stickprov av slipers av den andra typen. Vid kontroll av bärighet i de två stickproven har man fått fram följande information:

typ	$\bar{x}$	s	n
1	50.60	5.16	12
2	52.50	3.97	16

   (a) Testa $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$ på 5% risknivå.

   (b) På den aktuella delen av järnvägsnätet är en fjärdedel av alla slipers av den första typen medan resterande är av den andra typen. Medelvärdet över hela den här delen av nätet ska skattas. Beräkna ett 95% konfidensintervall för $0.25\mu_1 + 0.75\mu_2$.

4. (4 poäng) Se bakgrund i tidigare uppgift. Vid en tidigare olycka på en annan del av järnvägsnätet utfördes en viss kontroll som ger bara ett mått på om en sliper är godkänd eller underkänd enligt vissa kriterier.

   (a) Tabellen nedan visar data från stickprover av två typer av slipers. Kolumnen med rubriken u visar det antal som klassats som underkända bland dem som ingår i stickprovet.

typ	n	u
1	240	35
2	280	56

   Testa på 5% risknivå $H_0 : \pi_1 = \pi_2$, $H_1 : \pi_1 \neq \pi_2$

   (b) På den här delen av järnvägsnätet finns även slipers av en tredje typ. Ett stickprov av sådana slipers har också undersökts och data för den typen presenteras här på samma sätt som i deluppgift a.

typ	n	u
3	180	47

   Testa nollhypotesen att andelen underkända är densamma för alla tre typer av slipers. Använd 5% risknivå.

5. (3 poäng) Se bakgrund i tidigare uppgift. En viss kontroll ska utföras på ett stickprov om 15 slipers. Kontrollen visar bara om slipern är godkänd eller underkänd. Frågan avser andel underkända på den del av järnvägsnätet som stickprovet dras från. Man vill testa $H_0 : \pi = 0.125$, $H_1 : \pi \geq 0.125$ på 5% risknivå.

   (a) Man observerar 3 underkända bland de 15 som undersöks. Genomför testet.

   (b) Hur stor styrka har testet om andelen underkända egentligen är 0.4? En ledtråd kan vara att dela upplösningen i två steg, där man först beräknar för vilka antal som man drar slutsatsen att H_0 ska förkastas, och sedan beräknar sannolikheten att få ett av dessa värden.