

Tentamen

Linköpings universitet, Institutionen för datavetenskap, Avdelningen för Statistik och maskininlärning

Kurskod och namn:	732G01, Grundläggande statistik
Datum och tid:	2016-06-10, kl. 8-12
Jourhavande lärare:	Isak Hietala
Tillåtna hjälpmedel:	Dubbelsidigt handskrivet A4 med anteckningar, tabellsamling utan anteckningar, samt valfri räknedosa
Betygsgränser:	Tentamen omfattar totalt 20p. G från 12p och VG från 16p. Siffrorna i uppgifterna är påhittade. Saknas någon siffra för att kunna lösa uppgiften, skriv då tydligt ut att du saknar denna information, anta ett godtyckligt värde och lös uppgiften med detta antagande.

Redovisa, tolka och motivera tydligt alla dina lösningar!

Uppgift 1 (3p)

Sannolikheten att en händelse, A, sker är 0.2. Sannolikheten att en annan händelse, B, sker är 0.8.

- a) Bedöm om händelserna är oberoende givet att sannolikheten för att båda händelserna inträffar samtidigt är 0.16. (2p)

Kontrollera uttrycket:

$$Pr(A) * Pr(B) = Pr(A \cap B)$$
$$0.2 * 0.8 = 0.16$$

Händelserna är oberoende.

- b) Visualisera informationen om händelserna i ett Venn-diagram. (1p)

Uppgift 2 (3p)

Spelet Minecraft består utav kvadratiska block på 1x1x1 m av olika typer. Sannolikheten att blocket är av typen diamant är ca 1.192 procent.

- a) Vad är sannolikheten att man på ett område omfattande 776 block hittar högst 8 diamantblock? (3p)

$$X \sim bin(n = 776; \pi = 0.01192)$$
$$X \approx po(\mu = n * \pi = 9.24992)$$
$$Pr(X \leq 8) = 0.423$$

Normalapproximation också godkänd lösning.

Uppgift 3 (6p)

	n	$\bar{x}$	s
The Paradise	31	15.9	3.68
Selfridges	31	18.6	0.56

Två butiker, The Paradise och Selfridges, tävlar om att ha den högsta dagsförsäljningen. Under en månads tid mäter de båda butikscheferna försäljningen och en sammanfattning av denna data visas i ovanstående tabell. Försäljningen är angiven i tusentals pund.

- a) Butikschefen från Selfridges påstår att denna data bevisar att hans butik säljer mer än The Paradise. Kontrollera detta påstående med någon lämplig metod och använd 5 procents signifikans. (4p)

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 < 0$$

Eftersom $n^* > 30$ approximeras t med z .

$$z_{test} = \frac{15.9 - 18.6}{\sqrt{\frac{3.68^2}{31} + \frac{0.56^2}{31}}}$$

$$z_{test} = \frac{-2.7}{0.669}$$

$$z_{test} = -4.04$$

$$z_{krit} = z_{\alpha} = -1.64$$

Teststatistikan hamnar i förkastelseområdet. H_0 kan förkastas.

- b) Beräkna ett 95-procentigt konfidensintervall över The Paradise dagsförsäljning. Glöm inte att tolka resultatet! (2p)

Eftersom $n > 30$ approximeras t med z .

$$15.9 \pm 1.96 * \frac{3.68}{\sqrt{31}}$$

$$15.9 \pm 1.295$$

$$14.605 \leq \mu \leq 17.195$$

Uppgift 4 (5p)

Höjden, i cm, hos 10 stycken slumpmässigt utvalda blommor i Herr Blommenhofs trädgård är som följer :

32.1 30.2 29.8 29.3 25.1 28.9 29.2 28.1 34.7 27.7

- a) Beräkna medelvärdet och standardavvikelsen av höjden på dessa blommor. (2p)

$$\bar{x} = \frac{\sum x}{n} = \frac{295.1}{10} = 29.51$$

$$s_x = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} = \sqrt{\frac{8767.83 - \frac{295.1^2}{10}}{9}} = 2.57$$

- b) I Herr Blommenhofs trädgård finns det totalt 50 st blommor. Beräkna ett 99-procentigt konfidensintervall över den sanna medelhöjden av blommorna. (3p)

Eftersom $n < 30$ måste t -fördelningen användas.

$$\bar{x} - t_{n-1;1-\alpha} * \sqrt{\frac{s_x^2}{n} \left(1 - \frac{n}{N}\right)}$$

$$29.51 \pm 3.25 * \sqrt{\frac{2.57^2}{10} \left(1 - \frac{10}{50}\right)}$$

$$29.51 \pm 2.362$$

$$27.148 \leq \mu \leq 31.872$$

Uppgfit 5 (3p)

En stig delar sig i två, väg A och B. Bland 16 personer registrerades följande val av stig bland män och kvinnor.

	Man	Kvinna
A	5	3
B	1	7

- a) Vilka krav måste uppfyllas för att ett χ^2 -test ska kunna genomföras? (1p) Se kapitel 9 i kursboken.
b) Kan ett χ^2 -test beräknas på detta data?

	Man	Kvinna
A	3.00	5.00
B	3.00	5.00

Table 2: Förväntade frekvenser (E)

Kraven uppfylls inte och ett χ^2 -test kan ej beräknas