

Tentamen

Linköpings universitet, Institutionen för datavetenskap, STIMA

Kurskod och namn:	732G30, Grunder i statistisk metodik
Datum och tid:	2018-10-30, 8-12
Jourhavande lärare:	Isak Hietala
Tillåtna hjälpmedel:	Räknedosa av valfri modell, två stycken dubbelsidiga A4 med egna anteckningar, tabellsamling utan anteckningar
Betygsgränser:	Tentamen omfattar totalt 20 poäng, G från 12p, 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 (4p)

Ett slumpmässigt urval av 8 studenter på en idrottsutbildning har hoppat höjdhopp. Följande höjder i centimeter uppmättes: 179, 185, 182, 178, 194, 231, 183, 185.

- Är variabeln "höjd" kvalitativ eller kvantitativ? Vilken variabelskala är variabeln mätt på? (1p)
- Beräkna medelvärde och median för "höjd". Är det någon av dessa lägesmått som bättre beskriver mitten på datamaterialet jämfört med den andra? Motivera. (2p)
- Beräkna standardavvikelsen för variabeln. (1p)

Uppgift 2 (4p)

I rollspelet Dungeons and Dragons, kastar spelare en jämviktad 20-sidig tärning varje gång man ska utföra någonting, t.ex. öppna en låst dörr eller klättra upp för ett träd, för att se huruvida utförandet lyckas eller misslyckas. Ett lyckat utfall brukar generellt definieras som värden på minst 12.

- Vad är sannolikheten att man på 5 kast lyckas med det man ska utföra fler än 1 gång? (2p)

Om man lyckas kasta en 20:a på tärningen definieras det som ett perfekt utförande och man får någon form av bonus. Vi kan anta att man under ett spel kastar tärningen ca 250 gånger.

- Vad är sannolikheten att man under ett spel kastar fler än 20 bonusgivande kast? (2p)

Uppgift 3 (2p)

Tabell 1 visar en föreslagen sannolikhetsfördelning som ska ha beräknats enligt funktionen $P(X = x) = \frac{x}{\sum x}$.

Tabell 1: Föreslagen sannolikhetsfördelning

x	p(x)
2	0.050
5	0.125
8	0.200
11	0.275
14	0.350

Beräkna väntevärdet och variansen av den givna sannolikhetsfördelningen. (2p)

Uppgift 4 (6p)

6 stycken slumpmässigt utvalda företag inom försäkringsbranschen deltog i en undersökning om IT-kostnader:

Tabell 2: IT-kostnader i försäkringsbranschen

Företag	IT-kostnad i 1000 SEK
1	55
2	56
3	22
4	21
5	39
6	36

- Beräkna ett 95% konfidensintervall för IT-kostnader inom försäkringsbranschen. (1.5p)
- Den genomsnittliga IT-kostnaden oavsett bransch ligger på 35 000 SEK, testa på 10% signifikansnivå om försäkringsbranschen skiljer mot det genomsnittet. (3p)
- En motsvarande undersökning inom transportbranschen på 9 slumpmässigt utvalda företag hade en genomsnittlig IT-kostnad på 17 000 SEK med en varians 25 000. Beräkna ett 99% konfidensintervall för skillnaden i genomsnittlig IT-kostnad mellan försäkrings- och transportbranschen. (1.5p)

Uppgift 5 (4p)

Enligt Arbetsförmedlingen var arbetslösheten 8.5% i Östergötlands län och 8.3% i Kalmar län 2017.

- Testa på 5% signifikansnivå om arbetslösheten är större i Östergötlands län (458 150 invånare) mot Kalmar län (243 718 invånare). (3p)
- På vilken signifikansnivå skulle resultatet i a) kunna förkastas på om hypotesen var tvåsidig, rita upp och visa. (1p)