

TENTAMEN I TILLÄMPAD MATEMATIK OCH STATISTIK FÖR IT-FORENSIK. DEL 2: STATISTIK

7.5 HP

18 januari, 2019

Maxpoäng: 30p. **Betygsgränser:** 12p: betyg 3, 18p: betyg 4, 24p: betyg 5.

Hjälpmedel: Miniräknare och formelsamling. **Kursansvarig:** Eric Järpe, telefon 0729-77 36 26, 035-16 76 53.

Alla svar skall ges med 4 decimalers noggrannhet där ej annat anges. Till uppgifterna skall *fullständiga lösningar* lämnas. Lösningarna ska vara *utförligt* redovisade! Varje lösning ska börja överst på nytt papper. Endast en lösning per blad. Lösningar kommer finnas på internet: <http://dixon.hh.se/erja/teach> → Matematik och statistik för IT-forensik.

1. Bilinda observerar bilar som passerar ett övergångsställe mellan kl 10 och 11 varje dag under 2 veckor vilket ger följande statistik:

<i>Antal bilar per dag</i>	0	1	2	3	4	5+
<i>Antal dagar</i>	3	2	6	1	0	2

- (a) [2:1] Beräkna första kvartilen för variabeln $X =$ antal bilar per dag. (2p)
(b) [2:3] Ett vanligt antagande är att antalet bilar som passerar ett ställe är Poissonfördelat. Gör ett hypotestest på 5% signifikansnivå av om antagandet inte stämmer i detta fall. (4p)

2. [2:1] De länder som har störst utsläpp av växthusgaser sammanfattas med statistiken¹ (där utsläppssiffrorna är angivna i miljarder ton):

<i>Land</i>	Kina	USA	EU	Indien	Ryssland	Japan	Brasilien
<i>År 1990</i>	3.9	6.0	5.6	1.4	3.0	1.3	0.7
<i>År 2016</i>	13.0	6.4	4.4	3.7	2.2	1.4	1.2

Bilda med utläppsvärdena från *År 1990* som kovariat och utläppsvärdena från *År 2016* som respons en linjär modell för sambandet mellan dessa årtal. Vad blir interceptet och regressionskoefficienten i modellen? (4p)

3. [2:1] En skidskytte träffar sitt mål med sannolikhet 0.8 per skott vid liggande skytte och med sannolikhet 0.6 vid stående skytte. Under en tävling skjuter varje tävlande 2 gånger liggande och 2 gånger stående och för varje gång skjuts 3 skott. Vad är sannolikheten att ett skidskytt missar högst 2 skott under en tävling om man antar att skotten är oberoende av varandra? (4p)

¹Dessa data kommer från ett reportage i Hallandsposten 2018-12-03 om klimatmötet i Katowice i Polen.

4. [2:2] Gösta läser ur sin 408-sidiga bok varje kväll innan han somnar. Antalet läser sidor han läser per kväll har då väntevärdet 14 och variansen 36. Om han börjar läsa boken den 1 januari, vad är approximativt sannolikheten att han hinner läsa ut den innan februari? (3p)

5. Antag att $N(\mu, \sigma^2)$ och beräkna

(a) [2:2] $P(0.32 \leq X \leq 1.29)$ om $\mu = 0$ och $\sigma^2 = 1$. (2p)

(b) [2:2] $P(X \leq 1.29 | X \geq 0.32)$ om $\mu = \sigma^2 = 1$. (3p)

6. Grünolf har gjort en algoritm för numerisk lösning av polynomekvationer. Han mäter CPU-tiden i ms (millisekunder) för sin algoritm på 126 ekvationer vilket ger honom tiderna x_1, x_2, \dots, x_{126} där

$$\sum_{i=1}^{126} x_i = 1\,222 \text{ och } \sum_{i=1}^{126} x_i^2 = 12\,002.$$

Han klockar även Mathematicas CPU-tider för lösning av samma ekvationer och får då tiderna y_1, y_2, \dots, y_{126} där

$$\sum_{i=1}^{126} y_i = 1\,159 \text{ och } \sum_{i=1}^{126} y_i^2 = 10\,786.$$

(a) [2:2] Bilda ett 95% konfidensintervall för förväntad CPU-tid för Mathematicas algoritm under förutsättningen att standardavvikelsen är 1 ms. (2p)

(b) [2:3] Kan han bevisa att hans algoritm i genomsnitt exekverar snabbare än drömgränsen 10 ms på 5% signifikansnivå under antagandet att standardavvikelsen är 1 ms? Beräkna även p -värdet. (3p)

(c) [2:3] Antag nu att standardavvikelsen är okänd. Grünolf undrar om hans metod är på samma nivå som Mathematicas i bemärkelsen att han inte kan förkasta nollhypotesen i ett test på lämplig signifikansnivå av om den förväntade CPU-tiden för hans metod skiljer sig från Mathematicas. Kan han det? (3p)

LYCKA TILL!