

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

7.5 HP

januari, 2021

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

Hjälpmedel: Valfri miniräknare samt formelsamling som medföljer tentamens-texten.

Kursansvarig: Eric Järpe, telefon 0729-77 36 26.

Till uppgifterna skall *fullständiga lösningar* lämnas. Lösningarna ska vara *utförligt* redovisade! Bladen ska lämnas in i rätt ordning. Svara alltid med 4 decimalers noggrannhet om ej annat anges. Lösningar kommer finnas på internet: <http://dixon.hh.se/erja/teach> → Matematik och statistik för IT-forensik.

1. Henrik laddar upp filmer till sitt konto på Google drive. Utrymmet på hans konto är 15 GB och storleken av de filmer som Henrik laddar upp är fördelade $N(910, 50\,700)$ MB.
 - (a) [2:2] Vad är sannolikheten att en film är större än 1 GB? (2p)
 - (b) [2:2] Vad är sannolikheten att Henrik får plats med minst 18 filmer?(3p)
 - (c) [2:3] Antag att $\mu = E(\text{filmstorlek})$ är okänd men $\sigma^2 = V(\text{filmstorlek}) = 50\,700$ är känd. Hur många filmer måste observeras för att ett 95% konfidensintervall för μ ska bli högst 100 MB brett? (3p)

Henrik väljer slumpmässigt 12 filmer bland de 100-tals som han sparar på olika hårddiskar och observerar att de är

1 287 820 831 522 188 525 687 825 508 818 1 099 752

MB stora.

- (d) [2:1] Beräkna alla outliers i detta stickprov. (3p)
 - (e) [2:1] Skatta sannolikheten för att en film är mellan 800 och 900 MB.(2p)
 - (f) [2:3] Kan Henrik bevisa att egenskaperna *Filmstorleken är >800 MB* och *filmstorlekarna är ett jämnt antal MB* är oberoende? Gör ett test på 5% signifikansnivå och beräkna p -värdet. (4p)
2. [2:2] Ett visst sorts lösenord måste bestå av exakt 5 tecken, bara små bokstäver a–z och alla bokstäver måste vara olika. Hur många lösenord av denna sort kan konstrueras? (2p)
 3. [2:3] Ett företag har ett intranät där antalet intrångsförsök per månad observeras vara 23, 19, 22, 25 och 30. Kan man med dessa observationer bevisa på 5% signifikansnivå att det förväntade antalet intrång per månad är större än 20? (3p)

4. Holger ska sätta upp 70 fågelholkar i skogen och notera om de befåglas (dvs befolkas av fåglar).

- (a) [2:2] Om antalet fåglar i en slumpmässigt vald holk är Poissonfördelat med $\lambda = 2.3$, vad är den betingade sannolikheten att 3 fågelholkar får åtminstone 3 fåglar givet att de får högst 5 fåglar? (3p)

För de 58 fågelholkar som fåglar flyttar in i noterar Holger trädets höjd, h_i , och den tid, t_i , det tar från uppsättandet till inflyttningen och beräknar:

$$\sum_{i=1}^{58} h_i = 515.2 \quad \sum_{i=1}^{58} h_i^2 = 5446.5 \quad \sum_{i=1}^{58} t_i = 1126.3 \quad \sum_{i=1}^{58} t_i^2 = 25411.8 \quad \sum_{i=1}^{58} h_i t_i = 8367.8$$

- (b) [2:1] Beräkna den linjära modellen med $T = \text{tid till inflyttning}$ som funktion av $H = \text{trädets höjd}$. (3p)
- (c) [2:1] Skatta korrelation mellan variablerna T och H . (2p)

LYCKA TILL!