

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

7.5 HP

juni, 2023

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

Hjälpmedel: Valfri miniräknare samt formelsamling som medföljer tentamenstexten.

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. Under åren 2017–2021 rapporterades följande statistik beträffande *Bedrägerier* och *Dataintrång*¹:

	2017	2018	2019	2020	2021
<i>Bedrägerier</i>	581	603	619	614	768
<i>Dataintrång</i>	73	68	76	82	118

(a) [2:1] Beräkna medianen för variabeln *Bedrägerier*. (2p)

(b) [2:1] Beräkna förklaringsgraden för en linjär modell av *Bedrägerier* som funktion av *Dataintrång*. (4p)

Antag nu att *Bedrägerier* och *Dataintrång* är normalfördelade variabler och gör ett

(c) [2:3] hypotestest på 5% signifikansnivå av om det förväntade antalet *Bedrägerier* är mer än dubbelt så stort som det förväntade antalet *Dataintrång* per år. (4p)

(d) [2:3] 95% konfidensintervall för variabeln *Dataintrång* under det extra antagandet att dess varians är känd = 100. (2p)

2. [2:1] Vid en högriskmatch i fotboll gör 10 personer sig skyldiga till brott. Vid ett tillslag grips 15 personer ur en grupp om 20 personer som polisen bedömer som misstänkta och där samtliga de 10 skyldiga personerna befins. Vad är sannolikheten att man har gripit åtminstone 8 av de skyldiga huliganerna?(4p)

3. Antag att $X \in Poi(3.5)$. Vad är då

(a) [2:2] $P(X = 3)$? (1p)

(b) [2:2] $P(X > 3 | X \leq 5)$? (3p)

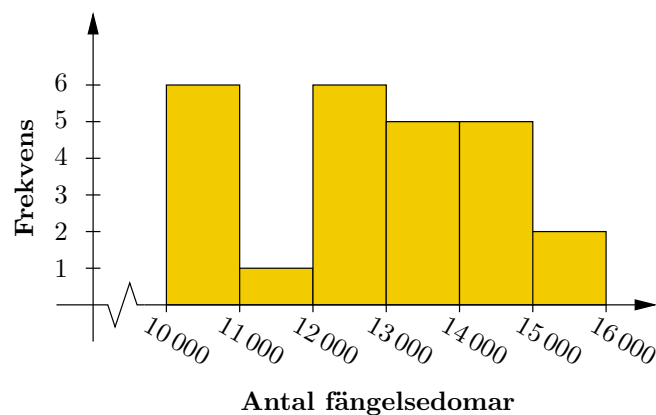
¹Källa: Brottsförebyggande Rådet, URL:

<https://www.bra.se/statistik/kriminalstatistik/personer-lagforda-for-brott.html>.

4. [2:2] Om $X \in N(e, e^2)$, vad är $P(e^X < 7e^e)$? (3p)

5. [2:2] Cineasten Filimon väljer ut 89 filmer till sin filmfestival. Denna filmfestival är speciell på så vis att alla filmer ska visas i ett sträck efter varandra: ingen film mer än en gång, inga pauser och ingen film visas samtidigt som någon annan. Om varje film har 2 timmars förväntad speltid och varje films speltid har varians 0.5, vad är approximativt sannolikheten att festivalen kommer att vara en hel vecka? (3p)

6. [2:3] Fördelningen av fängelsedomar utdelade i Sverige per år under de 25 åren 1995–2019 är redovisad i histogrammet nedan¹.



Kan man med dessa data bevisa att det förväntade antalet observationer *inte* minskar med 1 från intervall till intervall (dvs fördelningen under H_0 är sådan att det förväntade antalet observationer på intervallet $[10\,000, 11\,000)$ är 1 mer än det förväntade antalet observationer på intervallet $[11\,000, 12\,000)$, att förväntat antal observationer på intervallet $[11\,000, 12\,000)$ är 1 mer än förväntat antal observationer på intervallet $[12\,000, 13\,000)$, osv)? Gör ett hypotestest på 5% signifikansnivå. (4p)

LYCKA TILL!