

# TENTAMEN I MATEMATISK STATISTIK FÖR TILLÄMPAD ARTIFICIELL INTELLIGENS

## 7.5 HP

oktober, 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> → Matematisk statistik.

1. Ett viss AI tar olika lång tid på sig att exekvera, bl.a. beroende på hur mycket data som ska processas. Följande tider för hela processen, från inmatning av data till färdig klassificering, har observerats

<i>Storlek av data (GB)</i>	0.25	0.31	1.38	0.59
<i>Tid för processen (minuter)</i>	1.5	1.7	2.8	1.9

Beräkna

- (a) stickprovsstandardavvikelsen för variabeln *Storlek av data*. (3p)
- (b) intercept och regressionskoefficient i en linjär modell med variabeln *Tid för processen* som funktion av variabeln *Storlek av data*. (3p)
2. Då Henrik kastar pil missar han sitt mål med sannolikhet 81%. Dessutom är det 27% sannolikhet att någon börjar prata med honom då han står och siktar och då missar han sitt mål med 39% sannolikhet. Vad är den betingade sannolikheten att någon har stört Henrik om man får reda på att han missade målet? (3p)
3. En drickaback rymmer 20 flaskor. På ett lager fylls backar med läsk och öl. Där finns 30% läskflaskor och resten ölflaskor och man fyller backarna slumpmässigt med de två sorterna. Vad är sannolikheten
- (a) att en back innehåller minst 5 läskflaskor? (2p)
- (b) approximativt att en back i genomsnitt innehåller minst 6.2 läskflaskor om man fyller 37 backar? (3p)
4. Antag att  $X \in N(3, 2)$ . Beräkna
- (a)  $P(X \leq 3.2)$ . (2p)
- (b)  $P(\sqrt{X^2 + 1} \leq 3 \mid X > 1)$ . (3p)

5. Det är visserligen flera månader kvar till jul men planeringen börjar ovanligt tidigt i år. Tindra, Nicholas och Sirius ska vara lucia, tomte respektive stjärngosse i årets luciatåg på förskolan Sagolandet.

- (a) Vad blir styrkan av ett hypotestest med 5% signifikansnivå av om sannolikheten för att det är minst 5 tomtar är större än 50% i ett tåg med 15 barn då sannolikheten i själva verket är 75%? (4p)

De senaste fem åren har det varit

<i>År</i>	'18	'19	'20	'21	'22
<i>Antal lucior</i>	3	5	2	4	3
<i>Antal tomtar</i>	4	7	4	5	6
<i>Total antal barn i tåget</i>	17	14	13	13	17

- (b) Bilda ett 95% konfidensinterall för det förväntade antalet lucior per luciatåg. (3p)

Med statistik från åren 2008–2022 befinns antalet stjärngossar ha varit

<i>År</i>	'08	'09	'10	'11	'12	'13	'14	'15	'16	'17	'18	'19	'20	'21	'22
<i>Antal stjärngossar</i>	1	5	0	0	2	2	3	0	3	2	1	2	2	0	2

- (c) Avgör om man på 5% signifikansnivå kan bevisa att antalet stjärngossar per år ej är Poissonfördelat. Vad kan man säga om  $p$ -värdet för testet? (4p)

*LYCKA TILL!*