

This exam paper is available in both Danish and English. The English version appears after the Danish version.

Opgavesæt:

Skriftlig prøve: 28.05.2026

Kursusnavn og -nr.: 02402 Statistik (Polyteknisk grundlag)

Varighed: 4 timer

Tilladte hjælpemidler: Alle skriftlige hjælpemidler samt lommeregner af typen TI30XS eller TI30XB

De endelige svar skal afleveres ved udfyldelse af det særskilte “Answer Sheet”.

Opgavesættet består af 30 multiple choice-spørgsmål fordelt på 11 opgaver. **Kun “Answer Sheet” skal afleveres; selve opgavesættet må ikke afleveres.**

Multiple choice-spørgsmål: *Der er i hvert spørgsmål én og kun én korrekt svarmulighed. Endvidere er det ikke givet, at alle de angivne svarmuligheder er meningsfulde. Ved beregninger skal du altid afrunde dit resultat til det antal decimaler, der er angivet i svarmulighederne, før du vælger dit svar.*

Brug af Python til denne eksamen: *Denne eksamen indeholder Python-kode. Bemærk, at vi anvender følgende biblioteker og forkortelser:*

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import scipy.stats as stats
import statsmodels.api as sm
import statsmodels.formula.api as smf
import statsmodels.stats.power as smp
import statsmodels.stats.proportion as smprop
```

Fortsæt på side 3

Opgave I

Lad X og Y være uafhængige stokastiske variable. Forventningsværdierne er $\mathbb{E}[X] = 5$ og $\mathbb{E}[Y] = -3$, mens varianserne er $\mathbb{V}[X] = 9$ og $\mathbb{V}[Y] = 16$.

Spørgsmål I.1 (1)

Hvad er kovariansen mellem X og Y ?

1 ☐ -25

2 ☐ -5

3 ☐ 0

4 ☐ 5

5 ☐ 25

6 ☐ Ved ikke / Intet svar

Spørgsmål I.2 (2)

Lad $Z = -4X + 3Y$. Hvad er forventningsværdien af Z ?

1 ☐ -29

2 ☐ -27

3 ☐ 0

4 ☐ 27

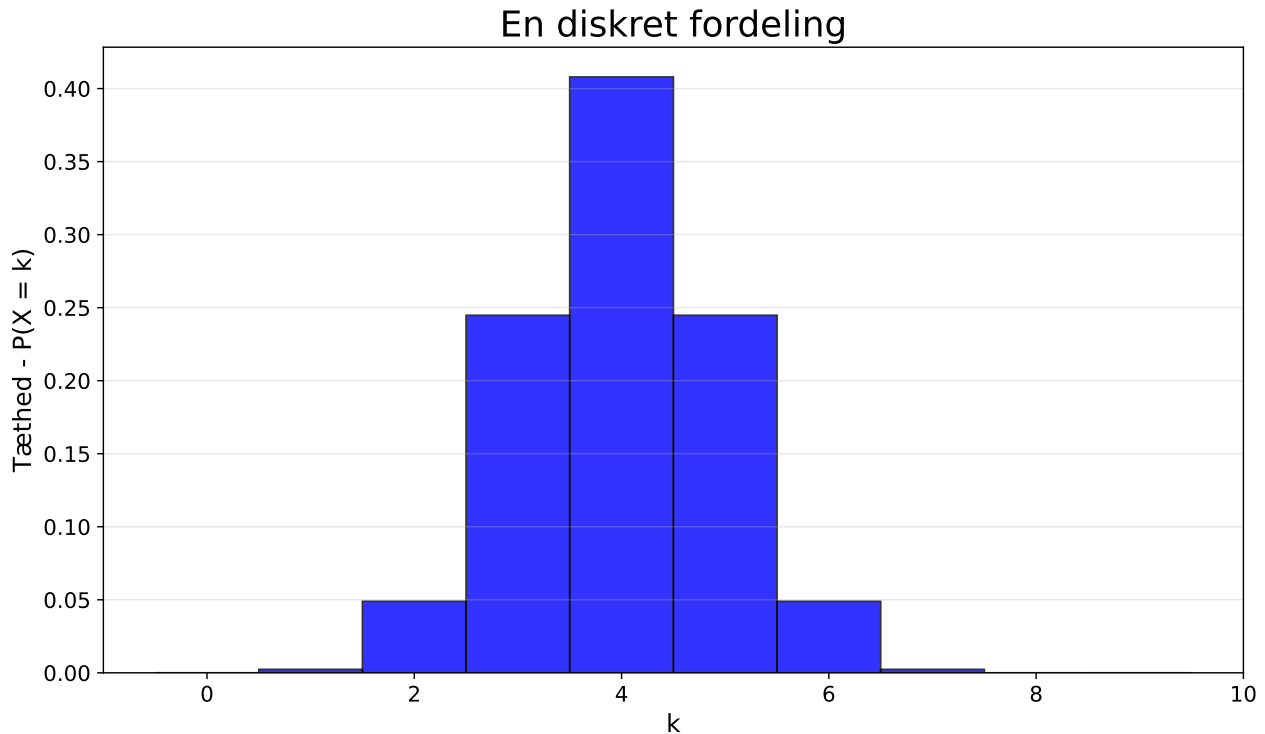
5 ☐ 29

6 ☐ Ved ikke / Intet svar

Fortsæt på side 4

Opgave II

Betragt den følgende diskrete fordeling på de ikke-negative hele tal repræsenteret ved dens tæthedsfunktion (pdf), og lad X følge denne fordeling.



Spørgsmål II.1 (3)

Hvilken af følgende muligheder angiver fordelingen korrekt?

- 1 ☐ Det er en binomialfordeling med $n = 10$ og $p = 0.5$.
- 2 ☐ Det er en binomialfordeling med $n = 10$ og $p = 0.9$.
- 3 ☐ Det er en hypergeometrisk fordeling med $n = 7$, $a = 8$ og $N = 14$.
- 4 ☐ Det er en Poissonfordeling med $\lambda = 4$.
- 5 ☐ Det er en Poissonfordeling med $\lambda = 10$.
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 5

Spørgsmål II.2 (4)

Hvad er $\mathbb{P}(X < 5)$?

- 1 ☐ Cirka 0.25
- 2 ☐ Cirka 0.30
- 3 ☐ Cirka 0.40
- 4 ☐ Cirka 0.70
- 5 ☐ Cirka 0.95
- 6 ☐ Ved ikke / Intet svar

Spørgsmål II.3 (5)

Hvad er variansen af X ?

- 1 ☐ Cirka 0.2
- 2 ☐ Cirka 0.7
- 3 ☐ Cirka 0.9
- 4 ☐ Cirka 4.2
- 5 ☐ Cirka 17.4
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 6

Opgave III

I et studie af strafudmåling indsamlede forskere data om fængselsstraffe (målt i år) idømt af forskellige dommere for sammenlignelige forbrydelser. Lad Y_{ij} betegne straf j givet af dommer i . Dataen modelleres som $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, hvor μ er den samlede gennemsnitlige straf på tværs af alle dommere, α_i er afvigelsen mellem den gennemsnitlige straf for dommer i og den samlede gennemsnitlige straf, og ε_{ij} er et stokastisk fejllid. Fejlledene antages at være uafhængige og ensfordelte med $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$, og alle straffe betragtes som uafhængige observationer.

Spørgsmål III.1 (6)

Hvilket udsagn er korrekt baseret på modellen beskrevet ovenfor?

- 1 ☐ Modellen tillader, at fejlleddene følger en vilkårlig kontinuert fordeling med middelværdi nul.
- 2 ☐ Modellen tillader, at dommere har forskellige gennemsnitlige straffe og forskellige varianser af straffene.
- 3 ☐ Modellen tillader, at dommere har forskellige gennemsnitlige straffe, men ikke forskellige varianser af straffene.
- 4 ☐ Modellen tillader, at dommere har forskellige varianser af straffene, men ikke forskellige gennemsnitlige straffe.
- 5 ☐ Modellen tillader hverken, at dommere har forskellige gennemsnitlige straffe eller forskellige varianser af straffene.
- 6 ☐ Ved ikke / Intet svar

Spørgsmål III.2 (7)

Forskerne forventer, at de forskellige dommeres gennemsnitlige straffe varierer meget, men at hver dommer giver meget konsistente straffe (dvs. at hver dommer har en tendens til at give ensartede straffe på tværs af sager). Hvis modelantagelserne er opfyldt, hvilket resultat vil da understøtte forskernes forventninger?

- 1 ☐ En lille værdi af $SS(\text{dommere})$ og en lille værdi af SSE .
- 2 ☐ En lille værdi af $SS(\text{dommere})$ og en stor værdi af SSE .
- 3 ☐ En stor værdi af $SS(\text{dommere})$ og en lille værdi af SSE .
- 4 ☐ En stor værdi af $SS(\text{dommere})$ og en stor værdi af SSE .
- 5 ☐ Intet resultat kan understøtte forskernes forventninger, hvis modelantagelserne er opfyldt.
- 6 ☐ Ved ikke / Intet svar

Modellen fittes til dataen inkluderet i studiet, som er vist i tabellen nedenfor:

Dommer A	Dommer B	Dommer C	Dommer D	Dommer E
4.3	7.8	10.1	4.7	13.2
14.1	8.4	10.5	10.3	17.9
6.9	10.2	10.4	8.1	12.4
20.5	11.3	9.9	7.6	12.5
2.2	9.3	10.1		10.1
10.8		11.0		19.8
		10.2		26.5
		12.6		11.1

Spørgsmål III.3 (8)

Hvilket udsagn om modelantagelserne er korrekt baseret på dataen? (Hint: prøv at visualisere dataen som boksplot.)

- 1 ☐ Modelantagelserne er opfyldt.
- 2 ☐ Modelantagelserne er ikke opfyldt, fordi dataen ikke er normalfordelt.
- 3 ☐ Modelantagelserne er ikke opfyldt, fordi studiet inkluderer færre end fem straffe fra en af dommerne.
- 4 ☐ Modelantagelserne er ikke opfyldt, fordi dommerne har meget forskellige varianser af straffene.
- 5 ☐ Modelantagelserne er ikke opfyldt, fordi dommerne har bidraget med forskellige antal straffe.
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 8

Opgave IV

Betragt en simpel lineær regressionsmodel:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2), \quad i = 1, \dots, n,$$

under de sædvanlige uafhængighedsantagelser.

Spørgsmål IV.1 (9)

Hvilket af følgende udsagn er forkert?

- 1 ☐ Modellen har tre parametre.
- 2 ☐ De stokastiske variable ε_i kaldes fejl.
- 3 ☐ Normalitetsantagelsen tjekkes med et normal QQ-plot af residualerne.
- 4 ☐ 95%-konfidensintervallet for hældningskoefficienten kan være bredere end 95%-konfidensintervallet for skæringspunktet med y-aksen.
- 5 ☐ 95%-konfidensintervallet for den forventede respons ved et givet x_0 kan være bredere end 95%-prædiktionsintervallet for en enkelt respons ved x_0 .
- 6 ☐ Ved ikke / Intet svar

Spørgsmål IV.2 (10)

Betragt en situation, hvor modellen er fittet til et datasæt med $n = 4$ observationer.

ID (i)	1	2	3	4
Observation (y_i)	4	6	8	10
Modelforudsigelse ($\hat{y}_i$)	4	5	10	9
Residual (e_i)	0	1	-2	1

Hvad er residualkvadratafgivelsessummen (RSS) for modellen?

- 1 ☐ Residualkvadratafgivelsessummen er -2.
- 2 ☐ Residualkvadratafgivelsessummen er 0.
- 3 ☐ Residualkvadratafgivelsessummen er 1.
- 4 ☐ Residualkvadratafgivelsessummen er 4.
- 5 ☐ Residualkvadratafgivelsessummen er 6.
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 9

Opgave V

Betragt en multipel kurvelineær regressionsmodel:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

hvor $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$, og den første række af designmatricen $\mathbf{X}$ er givet ved

$$[\cos(x_{11}) \quad \cos(x_{21}) \quad x_{11} + x_{21} \quad x_{11} \cdot x_{21}].$$

Spørgsmål V.1 (11)

Hvad er modelligningen for en enkelt observation?

- 1 ☐ $Y_i = \beta_0 + \beta_1 \cos(x_{1i}) + \beta_2 \cos(x_{2i}) + \beta_3 x_{1i} + \beta_4 x_{2i} + \beta_5 x_{1i} \cdot \beta_6 x_{2i} + \varepsilon_i$
- 2 ☐ $Y_i = \beta_0 + \beta_1 \cos(x_{1i}) + \beta_2 \cos(x_{2i}) + \beta_3(x_{1i} + x_{2i}) + \beta_4(x_{1i} \cdot x_{2i}) + \varepsilon_i$
- 3 ☐ $Y_i = \beta_0 + \beta_1 \cos(x_{1i}) + \beta_2 \cos(x_{2i}) + \beta_3 x_{1i} + \beta_4 x_{2i} + \beta_5(x_{1i} \cdot x_{2i}) + \varepsilon_i$
- 4 ☐ $Y_i = \beta_1 \cos(x_{1i}) + \beta_2 \cos(x_{2i}) + \beta_3 x_{1i} + \beta_4 x_{2i} + \beta_5(x_{1i} \cdot x_{2i})$
- 5 ☐ $Y_i = \beta_1 \cos(x_{1i}) + \beta_2 \cos(x_{2i}) + \beta_3(x_{1i} + x_{2i}) + \beta_4(x_{1i} \cdot x_{2i}) + \varepsilon_i$
- 6 ☐ Ved ikke / Intet svar

Spørgsmål V.2 (12)

Hvordan beregnes residualvektoren (vektoren med residualerne), $\mathbf{r}$?

- 1 ☐ $\mathbf{r} = \boldsymbol{\epsilon}$
- 2 ☐ $\mathbf{r} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$
- 3 ☐ $\mathbf{r} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$
- 4 ☐ $\mathbf{r} = \hat{\mathbf{Y}} - \mathbf{X}\hat{\boldsymbol{\beta}}$
- 5 ☐ $\mathbf{r} = \mathbf{Y} - (\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon})$
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 10

Modelselektionen viser, at den mest velegnede model er

$$Y_i = \beta_1 \cos(x_{1i}) + \beta_2(x_{1i} \cdot x_{2i}) + \varepsilon_i,$$

hvor fejlene, $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, antages uafhængige. Modellen fittes til et datasæt med 102 observationer ved brug af (ordinær) mindste kvadraters metode, hvilket fører til følgende parameterestimer:

$$\hat{\beta}_1 = 24, \quad \hat{\beta}_2 = 15, \quad \hat{\sigma} = 4.$$

Spørgsmål V.3 (13)

Hvad er residualkvadratafgivelsessummen, $\text{RSS}(\hat{\beta}_1, \hat{\beta}_2)$, for den fittede model?

- 1 ☐ 1600
- 2 ☐ 1584
- 3 ☐ 801
- 4 ☐ 400
- 5 ☐ 396
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 11

Opgave VI

Hver uge trækker det danske statslotteri 7 tilfældige tal uden tilbagelægning fra tallene 1 til 36 (begge inkluderet). En lottokupon består af 7 forskellige tal, og for at vinde hovedpræmien skal alle 7 tal på en kupon matche de trukne tal. (Bemærk: Rækkefølgen på tallene har ikke nogen betydning).

Spørgsmål VI.1 (14)

Hvad er chancen (sandsynligheden) for at vinde hovedpræmien med en enkelt kupon i lotteriet?

- 1 ☐ 1 ud af 78.364.164.096
- 2 ☐ 1 ud af 42.072.307.200
- 3 ☐ 1 ud af 2.176.782.336
- 4 ☐ 1 ud af 8.347.680
- 5 ☐ 1 ud af 5.040
- 6 ☐ Ved ikke / Intet svar

Opgave VII

Betragt to uafhængige stokastiske variable $X_1 \sim \mathcal{N}(0, 1)$ og $X_2 \sim \mathcal{N}(0, 1)$. En teoretisk model foreskriver, at $Y = f(X_1, X_2) = X_1^2 + X_2^2$.

Spørgsmål VII.1 (15)

Hvad er variansen af Y ifølge den ikke-lineære fejlafhobningslov?

- 1 ☐ Variansen af Y er 0.
- 2 ☐ Variansen af Y er 2.
- 3 ☐ Variansen af Y er 4.
- 4 ☐ Variansen af Y er 8.
- 5 ☐ Variansen af Y er 16.
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 12

Opgave VIII

Et universitet undersøger effekterne af AI platforme på studerendes eksamenspræstationer.

Spørgsmål VIII.1 (16)

I et stort kursus har en stikprøve på 15 studerende, som har brugt den anbefalede platform, en gennemsnitseksamensscore på 82 point med en standardafvigelse på 6.2 point. Under en normalfordelingsantagelse, hvad er 95%-konfidensintervallet for middeleksamensscoren for de studerende, som bruger den anbefalede platform? (Man kan bruge, at $t_{0.975}(14) = 2.145$.)

- 1 ☐ (78.6, 85.4)
- 2 ☐ (79.7, 84.3)
- 3 ☐ (80.0, 84.0)
- 4 ☐ (81.0, 83.0)
- 5 ☐ (77.5, 86.5)
- 6 ☐ Ved ikke / Intet svar

Spørgsmål VIII.2 (17)

I et andet stort kursus har en stikprøve på 12 studerende, som har brugt en alternativ platform, en gennemsnitsscore på 79 point (standardafvigelsen var 5 point) til eksamen. Universitetet anvender en tosidet t -test til at sammenligne middelscoren for de studerende, der bruger den alternative platform, med den historiske gennemsnitsscore for kurset på 75 point. Testen giver en p -værdi på 0.018 og et 95%-konfidensinterval på (75.82, 82.18). Hvad er den korrekte konklusion på testen ved 5% signifikansniveauet?

- 1 ☐ Middelscoren for studerende, der bruger den alternative platform, er signifikant lavere end den historiske gennemsnitsscore for kurset.
- 2 ☐ Middelscoren for studerende, der bruger den alternative platform, er ikke signifikant højere end den historiske gennemsnitsscore for kurset.
- 3 ☐ Middelscoren for studerende, der bruger den alternative platform, er ikke signifikant forskellig fra den historiske gennemsnitsscore for kurset.
- 4 ☐ Middelscoren for studerende, der bruger den alternative platform, er signifikant forskellig fra den historiske gennemsnitsscore for kurset.
- 5 ☐ Der er 1.8% sandsynlighed for, at middelscoren for studerende, der bruger den alternative platform, er signifikant højere end den historiske gennemsnitsscore for kurset.
- 6 ☐ Ved ikke / Intet svar

Universitetet har udviklet en AI tutor, designet til at hjælpe studerende med at studere og forberede sig til forelæsninger. For at undersøge om de studerende bruger mindre tid på at forberede sig til forelæsninger, når de bruger AI tutoren sammenlignet med traditionelle forberedelsesmetoder, har universitetet fulgt to grupper af studerende, der fulgte de samme forelæsninger: en lille gruppe på 50 studerende, der brugte AI tutoren, og en større gruppe på 300 studerende, der brugte traditionelle forberedelsesmetoder. For hver forelæsning målte universitet den gennemsnitlige forberedelsestid for begge grupper:

Forelæsning	1	2	3	4	...	13
Gennemsnitlig forberedelsestid med AI tutor (min.)	30	45	80	110	...	95
Gennemsnitlig forberedelsestid uden AI tutor (min.)	30	45	100	120	...	100

Universitetet noterer, at den forventede forberedelsestid varierer meget mellem forelæsningerne (jf. tabellen), og at de ulige gruppestørrelser kan føre til forskellige gruppevarianser.

Spørgsmål VIII.3 (18)

Hvilken af følgende muligheder repræsenterer den mest velegnede metode til at teste middelforskellen i forberedelsestid mellem de to grupper givet dette forsøgsdesign?

- 1 ☐ En Welch t -test for to uparrede stikprøver
- 2 ☐ En t -test for to uparrede stikprøver med sammenvejet varians
- 3 ☐ En parret t -test
- 4 ☐ En χ^2 -test
- 5 ☐ En ensidet variansanalysetest
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 14

Spørgsmål VIII.4 (19)

Universitetet beslutter at bruge en anden type test og vælger et signifikansniveau på 5%. For de givne parametre har testen en statistisk styrke på 90%. Givet at modelantagelserne er opfyldt, og at der er en reel forskel i middelforberedelsestiden mellem de to grupper, hvad er sandsynligheden for, at testen når frem til den korrekte konklusion?

- 1 ☐ 95%
- 2 ☐ 90%
- 3 ☐ 10%
- 4 ☐ 5%
- 5 ☐ 2.5%
- 6 ☐ Ved ikke / Intet svar

Spørgsmål VIII.5 (20)

I den alternative test beregnes nogle korrigerede forberedelsestider (som er strengt positive), som testen kræver skal være tilnærmelsesvis normalfordelte. Et histogram viser imidlertid, at de korrigerede forberedelsestider er højreskæve. Hvilken af følgende transformationer bør ikke overvejes, hvis målet er at transformere de korrigerede forberedelsestider, så de bliver tilnærmelsesvis normalfordelte?

- 1 ☐ Kubikrodstransformationen, dvs. at transformere x til $x^{1/3}$.
- 2 ☐ Den eksponentielle transformation, dvs. at transformere x til $\exp(x)$.
- 3 ☐ Logaritmetransformationen, dvs. at transformere x til $\log(x)$.
- 4 ☐ Reciproktransformationen, dvs. at transformere x til $1/x$.
- 5 ☐ Kvadratrodstransformationen, dvs. at transformere x til $\sqrt{x}$.
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 15

Opgave IX

Et universitet vil undersøge sammenhængen mellem studerendes brug af digitale træningsteknologier og deres eksamensscore. 240 studerende blev tilfældigt tildelt en af tre forskellige digitale træningsteknologier: Videobaseret Læring (VBL), Gamificeret Læringsplatform (GLP) og Interaktive Simuleringer (IS). Deres eksamensscore blev inddelt i tre niveauer: Under Middel, Middel og Over Middel.

Eksamensscore	VBL	GLP	IS	Rækkesum
Under Middel	18	12	10	40
Middel	32	26	22	80
Over Middel	30	38	52	120
Søjlesum	80	76	84	240

Studerende med Middel eller Over Middel eksamensscore betragtes som ”succesfulde elever”. Tabellen nedenfor med fraktiler fra standardnormalfordelingen skal bruges til at løse nogle af spørgsmålene i denne opgave.

Fraktil	$q_{0.90}$	$q_{0.95}$	$q_{0.975}$	$q_{0.99}$
Værdi	1.282	1.645	1.960	2.326

Spørgsmål IX.1 (21)

Under nulhypotesen, at fordelingen af eksamensscore er den samme på tværs af alle tre teknologier, hvad er så det forventede antal elever, som bruger Videobaseret Læring (VBL) med en Under Middel eksamensscore?

- 1 ☐ 12.90
- 2 ☐ 13.33
- 3 ☐ 14.00
- 4 ☐ 15.12
- 5 ☐ 18.50
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 16

Spørgsmål IX.2 (22)

Hvad er 95%-konfidensintervallet for den overordnede andel af succesfulde elever baseret på dataen?

- 1 ☐ [0.786, 0.880]
- 2 ☐ [0.701, 0.812]
- 3 ☐ [0.692, 0.833]
- 4 ☐ [0.688, 0.784]
- 5 ☐ [0.622, 0.754]
- 6 ☐ Ved ikke / Intet svar

Spørgsmål IX.3 (23)

Er der en signifikant forskel i andelen af succesfulde elever mellem VBL og IS grupperne på et 5% signifikansniveau? (Hint: beregn teststørrelsen under $H_0 : p_{VBL} - p_{IS} = 0$, hvor p_{VBL} er andelen af studerende i VBL-gruppen, der betragtes som succesfulde elever, medens p_{IS} er den tilsvarende andel inden for IS-gruppen.)

- 1 ☐ Der er ingen signifikant forskel, da den observerede teststørrelse $z_{\text{obs}} = -1.56 > -1.96$.
- 2 ☐ Der er ingen signifikant forskel, da den observerede teststørrelse $|z_{\text{obs}}| = |-2.12| > 1.96$.
- 3 ☐ Der er ingen signifikant forskel, da den observerede teststørrelse $z_{\text{obs}} = -1.80 > -1.96$.
- 4 ☐ Der er en signifikant forskel, da den observerede teststørrelse $z_{\text{obs}} = -2.27 < -1.96$.
- 5 ☐ Der er en signifikant forskel, da den observerede teststørrelse $z_{\text{obs}} = -2.13 < -1.96$.
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 17

Spørgsmål IX.4 (24)

Hvad er 95%-konfidensintervallet for forskellen i andelen af Over Middel eksamensscorer mellem IS og GLP grupperne?

- 1 ☐ $[0.113, 0.279]$
- 2 ☐ $[0.105, 0.293]$
- 3 ☐ $[0.091, 0.286]$
- 4 ☐ $[0.082, 0.312]$
- 5 ☐ $[-0.034, 0.272]$
- 6 ☐ Ved ikke / Intet svar

Spørgsmål IX.5 (25)

I den sædvanlige test af uafhængighed mellem typen af digital træningsteknologi og de studendes eksamensscore, hvilken fordeling følger teststørrelsen under nulhypotesen om uafhængighed?

- 1 ☐ En F -fordeling med 3 og 3 frihedsgrader
- 2 ☐ En F -fordeling med 2 og 2 frihedsgrader
- 3 ☐ En χ^2 -fordeling med 9 frihedsgrader
- 4 ☐ En χ^2 -fordeling med 6 frihedsgrader
- 5 ☐ En χ^2 -fordeling med 4 frihedsgrader
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 18

Opgave X

I en tosidet variansanalyse(ANOVA)model

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2),$$

hvor α_i repræsenterer behandlingseffekten og β_j repræsenterer blokeffekten, udføres en analyse for at undersøge signifikansen af begge faktorer.

Spørgsmål X.1 (26)

Hvilket af følgende udsagn er forkert?

- 1 ☐ Blokkene medtages for at reducere residualvariationen.
- 2 ☐ Behandlings- og blokeffekterne testes med separate F -tests med forskellige frihedsgrader.
- 3 ☐ Det antages, at fejlene i modellen er normalfordelte og uafhængige.
- 4 ☐ Blokkene medtages for at øge antallet af behandlingsgrupper.
- 5 ☐ Den totale variation kan opspaltes i behandlings-, blok- og residualkomponenter.
- 6 ☐ Ved ikke / Intet svar

Betragt variansanalyseskemaet (ANOVA-tabellen) fra analysen:

Kilde	Df	F-værdi	Pr(>F)
Behandling	2	74.40	0.0001
Blok	3	6.37	0.0270
Residualer	6	—	—

Fortsæt på side 19

Spørgsmål X.2 (27)

Hvilken af følgende fortolkninger er korrekt?

- 1 ☐ Både behandlings- og blokeffekterne er statistisk signifikante ved 1% signifikansniveauet.
- 2 ☐ Både behandlings- og blokeffekterne er statistisk signifikante ved 5% signifikansniveauet.
- 3 ☐ Blokeffekten er ikke statistisk signifikant ved 5% signifikansniveauet.
- 4 ☐ Residualerne er statistisk signifikante ved 5% signifikansniveauet, hvilket indikerer, at modellen er forkert.
- 5 ☐ Behandlingseffekten er ikke statistisk signifikant på ved 1% signifikansniveauet.
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 20

Opgave XI

I klassisk mekanik beskrives bevægelsen af et objekt under konstant acceleration af ligningen

$$v(t) = v_0 + at,$$

hvor $v(t)$ er hastigheden til tiden t , v_0 er starthastigheden (hastigheden til tiden $t = 0$) og a er den konstante acceleration.

I et eksperiment måles et sæt af hastigheder til forskellige tidspunkter for et objekt, der bevæger sig under konstant acceleration. Grundet måleusikkerheder er de observerede hastigheder dog forskellige fra den teoretiske model. For at håndtere dette, indføres et fejlelement, hvilket fører til den lineære regressionsmodel:

$$v_i = v_0 + at_i + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2),$$

hvor fejlene antages uafhængige.

Det følgende output er givet:

OLS Regression Results						
=====						
Dep. Variable:	v		R-squared:	1.000		
Model:	OLS		Adj. R-squared:	1.000		
No. Observations:	100		F-statistic:	3.071e+05		
Covariance Type:	nonrobust		Prob (F-statistic):	3.89e-173		
=====						
	coef	std err	t	P> t	[0.025	0.975]

Intercept	4.4695	0.421	10.629	0.000	3.635	5.304
t	4.0064	0.007	554.190	0.000	3.992	4.021
=====						

Spørgsmål XI.1 (28)

Hvad er de estimerede parameterverdier?

- 1 ☐ $\hat{v}_0 = 554.190$ og $\hat{a} = 10.629$
- 2 ☐ $\hat{v}_0 = 10.629$ og $\hat{a} = 554.190$
- 3 ☐ $\hat{v}_0 = 0.421$ og $\hat{a} = 0.007$
- 4 ☐ $\hat{v}_0 = 4.470$ og $\hat{a} = 4.006$
- 5 ☐ $\hat{v}_0 = 4.006$ og $\hat{a} = 4.470$
- 6 ☐ Ved ikke / Intet svar

Fortsæt på side 21

De følgende tal kan bruges i det næste spørgsmål:

```
print(np.round(stats.t.ppf(q = [0.995,0.99,0.975,0.95,0.90], df = 100-2),4))
```

[2.6269 2.365 1.9845 1.6606 1.2902]

Spørgsmål XI.2 (29)

Hvad er 99%-konfidensintervallet for v_0 ?

- 1 ☐ [3.364, 5.575]
- 2 ☐ [3.474, 5.465]
- 3 ☐ [3.635, 5.304]
- 4 ☐ [3.990, 4.023]
- 5 ☐ [3.992, 4.021]
- 6 ☐ Ved ikke / Intet svar

Spørgsmål XI.3 (30)

Betragt nulhypotesen $\mathcal{H}_0 : v_0 = 5$ og den tosidede modhypotese $\mathcal{H}_1 : v_0 \neq 5$. Hvad er den observerede teststørrelse (t_{obs}) under nulhypotesen?

- 1 ☐ -2.360
- 2 ☐ -1.260
- 3 ☐ 1.260
- 4 ☐ 10.629
- 5 ☐ 554.190
- 6 ☐ Ved ikke / Intet svar

Opgavesættet er slut.