

Skriftlig prøve: 15. december 2004

Kursus navn og nr: **Introduktion til Statistik, 02402**

Tilladte hjælpemidler: Alle

Dette sæt er besvaret af eksaminant

nr

Opgavesættet består af 30 spørgsmål af "multiple choice" typen. Bevarelsene af "multiple choice" spørgsmålene anføres ved at udfylde skemaet på forsiden (denne side), med numrene på de svarmuligheder, du mener er de korrekte.

Et forkert svar kan rettes ved at "sværte" det forkerte svar over og anføre det rigtige i stedet. Er der tvivl om meningen med en rettelse, eller er der anført flere end ét nummer ved et spørgsmål, betragtes spørgsmålet som ubesvaret. Kladde, mellemregninger eller andet tillægges **ingen** betydning, kun svarene i tabellen tæller.

Der gives 5 point for et korrekt "multiple choice" svar og -1 for et ukorrekt svar. Ubesvarede spørgsmål eller et 6-tal (svarende til "ved ikke") giver 0 point. Det antal point, der kræves for, at et sæt anses for tilfredsstillende besvaret, afgøres endeligt ved censureringen af sættene.

Den endelig besvarelse af opgaverne gøres ved at udfylde nedenstående skema. Kun skemaet tillægges betydning ved besvarelsen. Opgaveteksten skal dog afleveres i sin helhed, inden eksamen forlades – også selv om du vælger at aflevere blankt.

Opgave	I	II	III	IV	IV	IV	V	V	VI	VI
Spørgsmål	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Svar										

Opgave	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XV
Spørgsmål	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Svar										

Opgave	XV	XVI	XVII	XVII	XVIII	XVIII	XIX	XX	XXI	XXII
Spørgsmål	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
Svar										

Husk at forsyne opgavesættet med dit **nummer**. Sættets sidste side er nr 20; blad lige om og se, at den er der.

Fortsæt på side 2

Multiple choice opgaver Der gøres opmærksom på, at ideen med opgaverne er, at der er ét og kun ét rigtigt svar på de enkelte spørgsmål. Endvidere er det ikke givet, at alle de anførte alternative svarmuligheder er meningsfulde.

Opgave I

Spørgsmål I.1 (1): Man udtager tilfældigt to observationer fra en normalfordelt population med middelværdi $\mu = 3$ og varians $\sigma^2 = 2^2$, dvs. $X_1 \sim N(3, 2^2)$ og $X_2 \sim N(3, 2^2)$. Hvad bliver sandsynligheden for at summen af de to observationer overstiger 10?

- 1 ☐ $P(Z > \frac{4}{\sqrt{8}}) \approx 0.079$
- 2 ☐ $P(Z > -\frac{4}{\sqrt{8}}) \approx 0.921$
- 3 ☐ $2 \cdot P(Z < -\frac{4}{\sqrt{8}}) \approx 0.157$
- 4 ☐ $P(Z < -\frac{4}{8}) \approx 0.309$
- 5 ☐ $2 \cdot P(Z > \frac{4}{8}) \approx 0.617$
- 6 ☐ Ved ikke

Opgave II

I et forsøg undersøgtes effekten af tidlig kastrering af mus på forekomsten af diabetes. Ud af 100 mus af hankøn blev 50 tilfældig udvalgt og kastreret dagen efter fødslen. Efter 112 dage havde 26 af de kastrerede mus diabetes, mens kun 12 ud af de 50 ikke kastrerede mus havde diabetes.

Spørgsmål II.1 (2): Angiv et 99% konfidensinterval for forskellen mellem andelen af kastrerede mus, der får diabetes, og andelen af ikke-kastrerede mus, der får diabetes.

- 1 ☐ $0.28 \pm 1.96 \sqrt{\frac{0.52(1-0.52)}{100} + \frac{0.24(1-0.24)}{100}}$
- 2 ☐ $0.28 \pm 1.96 \sqrt{\frac{0.52(1-0.52)}{50} + \frac{0.24(1-0.24)}{50}}$
- 3 ☐ $0.28 \pm 2.576 \sqrt{\frac{0.52(1-0.52)}{100} + \frac{0.24(1-0.24)}{100}}$
- 4 ☐ $0.28 \pm 2.576 \sqrt{\frac{0.52(1-0.52)}{50} + \frac{0.24(1-0.24)}{50}}$
- 5 ☐ $0.28 \pm 2.576 \sqrt{\frac{0.52(1-0.48)}{50} + \frac{0.24(1-0.76)}{50}}$
- 6 ☐ Ved ikke

Fortsæt på side 3

Opgave III

Der planlægges et nyt forsøg med det formål at bestemme diabetesrisikoen for kastrede rotter, hvor man ikke på forhånd tør gætte på, hvilket niveau risikoen ligger på. Der ønskes en maximal fejl på 0.10 med 95% konfidens.

Spørgsmål III.1 (3): Hvor mange rotter skal undersøges?

- 1 ☐ $4 \cdot 1.96^2 / 0.1^2$
- 2 ☐ $1.96^2 / (4 \cdot 0.1^2)$
- 3 ☐ $4 \cdot 1.645^2 / 0.1^2$
- 4 ☐ $\frac{1}{4} \left(\frac{1.645}{0.1} \right)$
- 5 ☐ Ingen af ovenstående.
- 6 ☐ Ved ikke

Opgave IV

En operation i en produktionsvirksomhed kræver en vis oplæring for nyansatte. En ny oplæringsprocedure blev foreslået og sammenlignet med standard proceduren. Ni nyansatte oplærtes efter standard proceduren og ni andre oplærtes efter den nye procedure og derefter registreredes det hvor lang tid hver af de atten personer var om at gennemføre operationen. Målingerne blev (i minutter):

Ny procedure	35	31	29	25	34	40	27	32	31
Standard procedure	32	37	35	28	41	44	35	31	34

Spørgsmål IV.1 (4): Såfremt målingerne ikke kan antages at være normalfordelt, hvilken af følgende metoder er da mest hensigtsmæssig, hvis man vil sammenligne den nye procedure med standardproceduren?

- 1 ☐ En almindelig t-test
- 2 ☐ En tosidet variansanalyse
- 3 ☐ En parret t-test
- 4 ☐ En sign-test
- 5 ☐ En rank-sum test
- 6 ☐ Ved ikke

Fortsæt på side 4

Spørgsmål IV.2 (5): Såfremt målingerne for den nye procedure kan antages at være normalfordelt $N(\mu, \sigma^2)$, bliver den kritiske værdi for det sædvanlige test med $\alpha = 5\%$ for hypotesen $H_0: \mu = 30$ mod alternativet $H_1: \mu > 30$ (baseret på de 9 målinger)

- 1 ☐ $z_{0.025} = 1.960$
- 2 ☐ $z_{0.05} = 1.645$
- 3 ☐ $t(8)_{0.05} = 1.860$
- 4 ☐ $t(9)_{0.025} = 2.262$
- 5 ☐ $t(9)_{0.05} = 1.833$
- 6 ☐ Ved ikke

Spørgsmål IV.3 (6): Antag, at målingerne for den nye procedure kan antages at være normalfordelt $N(\mu, \sigma^2)$. Baseret på de 9 målinger oven for bliver estimatet for middelværdien $\hat{\mu} = 31.56$ og estimatet for spredningen $\hat{\sigma} = 4.475$. Et 95% konfidensinterval for spredningen bliver

- 1 ☐ $\frac{8 \cdot 4.475}{17.535} < \sigma < \frac{8 \cdot 4.475}{2.180}$
- 2 ☐ $\frac{3 \cdot 4.475}{2.180} < \sigma < \frac{3 \cdot 4.475}{17.535}$
- 3 ☐ $\frac{8 \cdot 4.475^2}{17.535} < \sigma^2 < \frac{8 \cdot 4.475^2}{2.180}$
- 4 ☐ $\frac{\sqrt{8} \cdot 4.475}{\sqrt{17.535}} < \sigma < \frac{\sqrt{8} \cdot 4.475}{\sqrt{2.180}}$
- 5 ☐ $31.56 \pm 2.306 \frac{4.475}{\sqrt{9}}$
- 6 ☐ Ved ikke

Fortsæt på side 5

Opgave V

Syv patienter med moderat forhøjet blodtryk, som havde lært at måle deres eget blodtryk, indgik i et forsøg for at undersøge deres blodtryk morgen og aften. Både morgen og aften målt blodtrykket både liggende og stående. De målte blodtryk (i mm HG) for de fire situationer er angivet i følgende tabel:

Patient	Morgen Liggende	Morgen Stående	Aften Liggende	Aften Stående
1	143	148	152	157
2	144	149	151	152
3	141	153	155	155
4	135	141	149	156
5	153	157	163	165
6	152	139	165	168
7	159	157	162	166

Spørgsmål V.1 (7): Hvilken type analyse er mest hensigtsmæssig med henblik på at udtale sig om, hvorvidt situationen betyder noget for blodtryksniveauet?

- 1 ☐ Test i en antalstabel
- 2 ☐ En parret t-test
- 3 ☐ En ensidet variansanalyse
- 4 ☐ En tosidet variansanalyse
- 5 ☐ Simpel lineær regression
- 6 ☐ Ved ikke

Spørgsmål V.2 (8): Hvad er den 3. kvartil (øvre) for den liggende morgenblodtryk?

- 1 ☐ 159
- 2 ☐ 152
- 3 ☐ 153
- 4 ☐ 152.5
- 5 ☐ 144
- 6 ☐ Ved ikke

Fortsæt på side 6

Opgave VI

I forbindelse med udviklingen af mulige lægemiddelstoffer designes på molekylært niveau en hel række kandidater, hvis egenskaber man prøver at udlede ud fra den molekylære struktur. For 832 stoffer sås følgende korrelationskoefficienter mellem de tre egenskaber MW (molekylær vægt), SURF (molekylært overfladeareal) og VOL (molekylært volumen):

	SURF	VOL	MW
SURF	1	0.995	0.949
VOL	0.995	1	0.952
MW	0.949	0.952	1

Spørgsmål VI.1 (9): Hvis man på disse data tilpasser følgende lineære regressionsmodel: $MW_i = \alpha + \beta \cdot SURF_i + \varepsilon_i$, hvor stor en del af variationen i MW_i bliver da forklaret?

- 1 ☐ 90.1%
- 2 ☐ 99.0%
- 3 ☐ 1.0%
- 4 ☐ 9.4%
- 5 ☐ Ingen af ovenstående
- 6 ☐ Ved ikke

Spørgsmål VI.2 (10): Antag at man på disse data har tilpasset følgende lineære regressionsmodel: $VOL_i = \alpha + \beta \cdot SURF_i + \varepsilon_i$. Hvis molekyle A har et overfladeareal (SURF) der er 200 arealenheder større end molekyle B (dvs. $SURF_A = SURF_B + 200$), så vil det forventede volumen af molekyle A være givet ved:

- 1 ☐ $VOL_A = 200 + VOL_B$
- 2 ☐ $VOL_A = 199 + VOL_B$
- 3 ☐ $VOL_A = 1 + VOL_B$
- 4 ☐ $VOL_A = VOL_B - 199$
- 5 ☐ Ingen af ovenstående
- 6 ☐ Ved ikke

Fortsæt på side 7

Opgave VII

I en undersøgelse af 712 lægemiddelkandidater fandt man en gennemsnitligt molekylvægt på 437 og en standardafvigelse på 117.

Spørgsmål VII.1 (11): Inden for hvilke grænser vil man forvente at 95% af denne type lægemiddelkandidater ligger?

- 1 ☐ $437 \pm 1.645 \frac{117}{437}$
- 2 ☐ $437 \pm 1.96 \frac{117^2}{437}$
- 3 ☐ $437 \pm 1.96 \cdot 117^2$
- 4 ☐ $437 \pm 1.96 \cdot 117$
- 5 ☐ $437 \pm 1.645 \cdot 117$
- 6 ☐ Ved ikke

Opgave VIII

I en undersøgelse testedes 712 lægemiddelkandidater for permeabilitet. De permeable stoffer (332 i alt) havde en gennemsnitlig molekylvægt på 373 og for de ikke-permeable (380 i alt) var den gennemsnitlig molekylvægt 492. Den sammenvejede spredning var $s_p = 100.7$. Det sædvanlige test for at påvise, at molekylvægten for permeable stoffer er mindre end for ikke-permeable stoffer bliver:

Spørgsmål VIII.1 (12):

- 1 ☐ $\frac{380-332}{\sqrt{100.7^2/380+100.7^2/332}}$
- 2 ☐ $\frac{373-492}{\sqrt{100.7^2/373+100.7^2/373}}$
- 3 ☐ $\frac{373-492}{\sqrt{100.7^2/332+100.7^2/380}}$
- 4 ☐ $\frac{380-332}{\sqrt{100.7^2/712+100.7^2/712}}$
- 5 ☐ $\frac{380-492}{\sqrt{100.7/332+100.7/380}}$
- 6 ☐ Ved ikke

Fortsæt på side 8

Opgave IX

I en undersøgelse testedes 712 lægemiddelkandidater for permeabilitet. En beregningsmodel forudså permeabilitetsegenskaben for 190 nye stoffer, som dernæst testedes i laboratoriet. Man fik følgende resultater

	Forudsagt ikke-permeabel	Forudsagt permeabel
Testet ikke-permeabel	77	18
Testet permeabel	7	88

Spørgsmål IX.1 (13): Kan man lave et χ^2 -test for at påvise, at beregningsmodellen er signifikant bedre end blot at gætte tilfældigt om et stof er permeabelt?

- 1 ☐ Nej, for ikke alle forventede antal er større end 15.
- 2 ☐ Ja, for alle forventede antal er større end 5.
- 3 ☐ Nej, for alle forventede antal er mindre end 5.
- 4 ☐ Ja, for de fleste forventede antal er større end 15.
- 5 ☐ Nej, for de fleste forventede antal er større end 15.
- 6 ☐ Ved ikke

Fortsæt på side 9

Opgave X

Fire smagsdommere vurderede sødheden på en skala mellem 0 og 10 for 5 produkter med følgende resultater:

	dom 1	dom 2	dom 3	dom 4
produkt 1	6	8	10	7
produkt 2	3	5	8	4
produkt 3	10	9	9	8
produkt 4	7	6	6	6
produkt 5	8	8	10	7

Spørgsmål X.1 (14): Den kritiske værdi for det sædvanlige test, hvor data antages at være normalfordelt, med $\alpha = 5\%$ for hypotesen om at der ikke er forskel sødhedsniveauerne i de fem produkter bliver:

- 1 ☐ $F(4, 15)_{0.05} = 3.06$
- 2 ☐ $F(4, 12)_{0.05} = 3.26$
- 3 ☐ $F(5, 12)_{0.05} = 3.11$
- 4 ☐ $F(3, 12)_{0.05} = 3.49$
- 5 ☐ $F(5, 15)_{0.05} = 2.90$
- 6 ☐ Ved ikke

Fortsæt på side 10

Opgave XI

Det såkaldte celletal i en mælkeprøve udtrykker graden af yverbetændelse. For mælkeprøver fra 166 malkefår (en enkelt prøve pr. får) fra i alt 3 besætninger målt logaritmen til celletallet. Gennemsnittet for alle 166 får var 4.896 I følgende tabel ses nogle basale beregningsstørrelser:

	Antal får	Gennemsnit	Spredning
Besætning 1	39	4.823	1.654
Besætning 2	67	5.293	1.706
Besætning 3	60	4.499	1.746

Spørgsmål XI.1 (15): Den sædvanlige teststørrelse for at undersøge om der er forskel på celletalsniveauerne i de tre besætninger bliver:

- 1 ☐ $\frac{[39 \cdot (4.823 - 4.896)^2 + 67 \cdot (5.293 - 4.896)^2 + 60 \cdot (4.499 - 4.896)^2]}{[39 \cdot 1.654^2 + 67 \cdot 1.706^2 + 60 \cdot 1.746^2]}$
- 2 ☐ $\frac{[38 \cdot 4.823^2 + 66 \cdot 5.293^2 + 59 \cdot 4.499^2]/2}{[38 \cdot 1.654^2 + 66 \cdot 1.706^2 + 59 \cdot 1.746^2]}$
- 3 ☐ $\frac{[39 \cdot (4.823 - 4.896)^2 + 67 \cdot (5.293 - 4.896)^2 + 60 \cdot (4.499 - 4.896)^2]/3}{[38 \cdot 1.654 + 66 \cdot 1.706 + 59 \cdot 1.746]/166}$
- 4 ☐ $\frac{[39 \cdot (4.823 - 4.896)^2 + 67 \cdot (5.293 - 4.896)^2 + 60 \cdot (4.499 - 4.896)^2]/2}{[38 \cdot 1.654 + 66 \cdot 1.706 + 59 \cdot 1.746]/163}$
- 5 ☐ $\frac{[39 \cdot (4.823 - 4.896)^2 + 67 \cdot (5.293 - 4.896)^2 + 60 \cdot (4.499 - 4.896)^2]/2}{[38 \cdot 1.654^2 + 66 \cdot 1.706^2 + 59 \cdot 1.746^2]/163}$
- 6 ☐ Ved ikke

Fortsæt på side 11

Opgave XII

I et klassisk studie observerede man ud af 6115 familier med 12 børn i alt 38100 drengefødsler ud af de i alt 73380 fødsler. Hvis binomialfordelingen skulle beskrive antallet af drengefødsler ud af 12 børn i en familie, hvor mange familier med præcis 6 drenge ville man så forvente?

Spørgsmål XII.1 (16):

- 1 ☐ $6115 \cdot 0.519^6(1 - 0.519)^6$
- 2 ☐ $73380 \cdot \binom{12}{6} 0.519^6(1 - 0.519)^6$
- 3 ☐ $\binom{12}{6} 0.519^6(1 - 0.519)^6$
- 4 ☐ $6115 \cdot \binom{12}{6} 0.519^6(1 - 0.519)^6$
- 5 ☐ $6115 \cdot \binom{12}{6} 0.5^6(1 - 0.5)^6$
- 6 ☐ Ved ikke

Opgave XIII

En butik har i gennemsnit 20 kunder i timen, som antages at ankomme helt tilfældigt. Man vil gerne belønne kunde nummer 100 hver dag. Kunde nummer 99 ankommer kl. 12.58.

Spørgsmål XIII.1 (17): Hvad er sandsynligheden for at kunde nummer 100 ankommer senest kl. 13.00?

- 1 ☐ e^{-6}
- 2 ☐ $1 - e^{-6}$
- 3 ☐ $1 - e^{-2/3}$
- 4 ☐ $e^{-3/2}$
- 5 ☐ $e^{-2/3}$
- 6 ☐ Ved ikke

Fortsæt på side 12

Opgave XIV

I en undersøgelse ved Lady Elliot Island, Australien, studeredes forekomsten af dyr på forskellige koraltyper. Der udtoges prøver fra tre forskellige koraltyper. I følgende tabel ses antallet af to forskellige sneglearter i disse prøver:

	Koraltype			Total
	I	II	III	
Art A	6	2	14	22
Art B	7	21	1	29
Total	13	23	15	51

Spørgsmål XIV.1 (18): Man får en χ^2 teststørrelse på 26.58. Kan man afvise, at der er samme fordeling af sneglearterne A og B på de tre koraltyper?

- 1 ☐ Ja, for P-værdien er 5%.
- 2 ☐ Ja, for P-værdien er 1%.
- 3 ☐ Nej, for P-værdien er større end 5%.
- 4 ☐ Nej, for P-værdien er mindre end 1%.
- 5 ☐ Ja, for P-værdien er mindre end 0.5%.
- 6 ☐ Ved ikke

Fortsæt på side 13

Opgave XV

Der udtages en tilfældig stikprøve på 14 fisk. For hver fisk målttes længden og tykkelsen. I følgende tabel ses de 14 længder og tykkelser (i mm):

Længde, x	Tykkelse, y
104.0	35.6
114.5	38.0
113.2	37.8
115.2	35.2
116.8	38.2
122.3	40.7
120.0	40.0
130.5	49.6
132.4	46.0
133.5	47.2
114.4	38.7
139.0	51.7
144.9	50.0
146.0	47.4
$\bar{x} = 124.8$	$\bar{y} = 42.6$
$s_x^2 = 168.3$	$s_y^2 = 33.5$
$S_{xy} = 898.16$	

Der arbejdes med en lineær model for tykkelsen som funktion af længden:

$$Y_i = \alpha + \beta x_i + \varepsilon_i$$

Spørgsmål XV.1 (19): Estimatet for hældningskoefficienten i denne model er givet ved:

1 ☐ $42.6 - \frac{898.16}{13 \cdot 168.3} \cdot 124.8$

2 ☐ $\frac{898.16/14}{13 \cdot 168.3} \cdot 124.8$

3 ☐ $42.6 - \frac{898.16}{68.3}$

4 ☐ $\frac{898.16}{13 \cdot 168.3}$

5 ☐ $\frac{898.16/14}{168.3}$

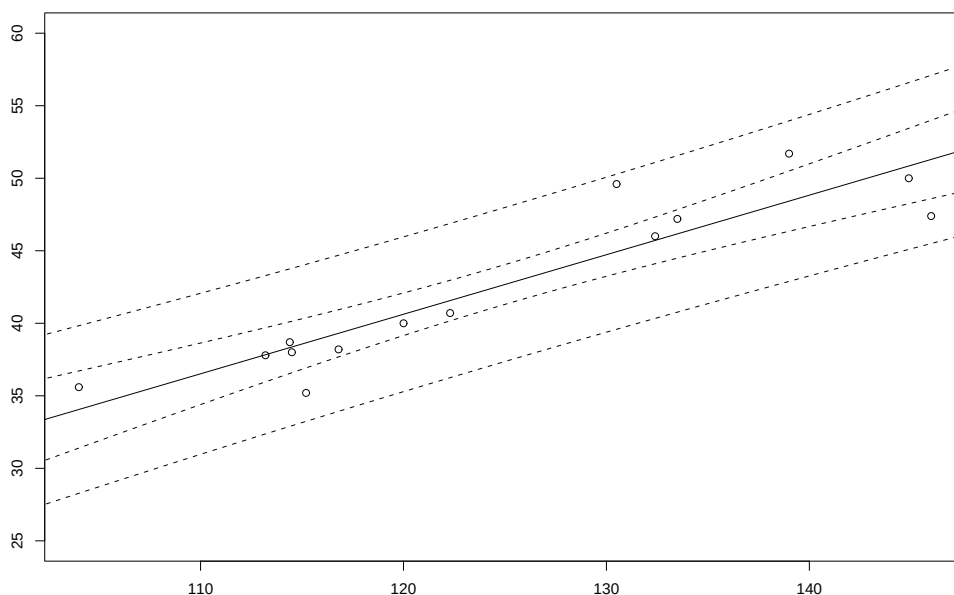
6 ☐ Ved ikke

Fortsæt på side 14

Spørgsmål XV.2 (20): Estimater for variansen for afvigelserne fra linien er givet ved

- 1 ☐ 33.5
- 2 ☐ $\frac{13 \cdot 33.5 - 898.16^2 / (13 \cdot 168.3)}{12}$
- 3 ☐ $13 \cdot 33.5 - 898.16^2 / (13 \cdot 168.3)$
- 4 ☐ $\sqrt{13 \cdot 33.5 - 898.16^2 / (13 \cdot 168.3)}$
- 5 ☐ $\frac{898.16^2 / (13 \cdot 168.3 - 13 \cdot 33.5)}{12}$
- 6 ☐ Ved ikke

I figuren ses de 14 samhørende værdier, den tilpassede linie samt 95% konfidensgrænserne for $\alpha + \beta x_i$ og 95% prædiktionsgrænserne.



Spørgsmål XV.3 (21): En ny fisk fanges, som har længden 110 mm. Et interval, der med 95% konfidens indeholder den rigtige tykkelse for denne fisk kan:

- 1 ☐ Aflæses på figuren til at være omtrent (34, 39)
- 2 ☐ Beregnes som $a + b \cdot 110 \pm 1.96\sqrt{33.5}$
- 3 ☐ Aflæses på figuren til at være omtrent (31, 42)
- 4 ☐ Beregnes som $b \cdot 110 \pm 1.96\sqrt{33.5}$
- 5 ☐ Aflæses på figuren til at være omtrent (27, 39)
- 6 ☐ Ved ikke

Fortsæt på side 15

Opgave XVI

I et frostlagringsforsøg med laks måltet indholdet af frie fedtsyrer i otte laks ved forsøgets start og igen efter 19 ugers lagring. Resultaterne ses i følgende tabel, der angiver procenten af frie fedtsyrer.

Laks	Start	Efter 19 uger
1	0.50	4.91
2	0.46	2.40
3	0.73	2.60
4	0.50	6.05
5	0.38	1.08
6	0.54	1.08
7	0.47	1.47
8	0.56	0.97

Spørgsmål XVI.1 (22): Hypotesen om ingen lagringseffekt mod alternativet at der sker en stigning i procenten af frie fedtsyrer testes med et sign test og $\alpha = 0.01$. Hvor mange stigninger skal der mindst til for at forkaste hypotesen?

- 1 ☐ 5
- 2 ☐ 6
- 3 ☐ 7
- 4 ☐ 8
- 5 ☐ 0
- 6 ☐ Ved ikke

Fortsæt på side 16

Opgave XVII

Ud af 170 vindere i ligeså mange hundevæddeløb havde de 35 haft inderbanen som startposition.

Spørgsmål XVII.1 (23): Et konfidensinterval for vindersandsynligheden p på inderbanen er:

$$0.1284 < p < 0.2833$$

Hvad er α ?

- 1 ☐ 5%
- 2 ☐ 10 %
- 3 ☐ 1.25%
- 4 ☐ 0.625%
- 5 ☐ 0.1%
- 6 ☐ Ved ikke

Spørgsmål XVII.2 (24): Antallet af vindere ud af de 170, der ikke havde haft inderbanen som startposition kunne have følgende fordeling:

- 1 ☐ t-fordeling med 1 frihedsgrad
- 2 ☐ Poissonfordeling med $\lambda = 35/170$
- 3 ☐ Binomialfordeling med $n = 170$ og $p = 135/170$
- 4 ☐ χ^2 -fordeling med 1 frihedsgrad
- 5 ☐ Poissonfordeling med $\lambda = 135/170$
- 6 ☐ Ved ikke

Fortsæt på side 17

Opgave XVIII

I en undersøgelse til vurdering af kvaliteten af oste indgik blandt andre tre oste, her betegnet nummer 1, 2 og 3. Vurderinger af bitterheden i de tre oste udført af 10 smagsdommere er vist i følgende tabel.

Dommer	Bittersmag		
	Ost 1	Ost 2	Ost 3
1	3.5	5.6	2.8
2	5.3	4.2	3.6
3	6.0	5.3	6.8
4	4.0	4.8	4.3
5	5.8	5.8	7.1
6	2.6	1.0	1.0
7	1.1	4.2	1.3
8	1.0	2.8	1.0
9	5.4	4.9	3.2
10	5.6	6.2	4.6

En passende analyse af disse data kan opgøres i følgende tabel:

Variations- kilde	Kvadratsum	Friheds- grader	Middel- kvadratsum	F	P-værdi
Oste	4.1407	2	2.0703	2.07	0.1555
Dommere	80.6120	9	8.9569	8.94	<0.0001
Fejl	18.0260	18	1.0014		

Spørgsmål XVIII.1 (25): Hvilket udsagn udtrykker mest korrekt de resultater vedrørende bittersmagen, der umiddelbart ses af denne tabel:

- 1 ☐ Der er forskel både på dommere og på oste.
- 2 ☐ Der er forskel på dommere men ikke på oste.
- 3 ☐ Der er forskel på oste men ikke på dommere.
- 4 ☐ Der er hverken forskel på dommere eller på oste.
- 5 ☐ Der er forskel både på dommere og på oste og ingen forskel på fejlen.
- 6 ☐ Ved ikke

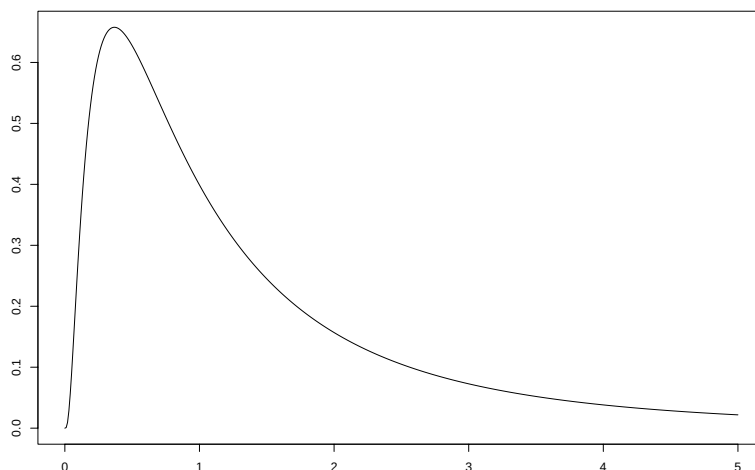
Fortsæt på side 18

Spørgsmål XVIII.2 (26): Hvis man fejlagtigt analyserede ovenstående data som en almindelig ensidig variansanalyse (fuldstændig randomiseret forsøg), så ville F-testet for hypotesen om ingen forskel på ostene blive:

- 1 ☐ $\frac{2.0703}{(80.6120+18.0260)/27}$
- 2 ☐ 2.07
- 3 ☐ 8.94
- 4 ☐ $\frac{2.0703}{8.9569}$
- 5 ☐ $\frac{8.9569}{(80.6120+18.0260)/27}$
- 6 ☐ Ved ikke

Opgave XIX

I nedestående figur ses en tæthedsfunktion:



Spørgsmål XIX.1 (27): Hvilken af følgende sandsynlighedsfordelinger er den eneste, der kunne være beskrevet ved ovenstående figur?

- 1 ☐ Eksponentialfordelingen.
- 2 ☐ Normalfordelingen.
- 3 ☐ Den uniforme fordeling.
- 4 ☐ Poissonfordelingen
- 5 ☐ Lognormalfordelingen
- 6 ☐ Ved ikke

Fortsæt på side 19

Opgave XX

Residualerne fra en lineær regressionsmodel kom i ordnet rækkefølge med følgende fortegn:

Observationsnummer	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Fortegn	+	-	-	+	+	+	-	-	+	-	+	-	+	+	-	-

Spørgsmål XX.1 (28): Den sædvanlige teststørrelse for at undersøge om residualernes fortegn udgør en tilfældig sekvens bliver:

1 ☐ $\frac{9 - (128/16 + 1)}{\sqrt{\frac{128(128 - 16)}{16^2 \cdot 15}}}$

2 ☐ $\frac{10 - 64/2}{\sqrt{\frac{64(1 + 16)}{12}}}$

3 ☐ $\frac{9 - 64/2}{\sqrt{\frac{64(1 + 16)}{12}}}$

4 ☐ $\frac{8/16 - \frac{1}{2}16}{\sqrt{\frac{8/16(1 - 8/16)}{16}}}$

5 ☐ $\frac{10 - (128/16 + 1)}{\sqrt{\frac{128(128 - 16)}{16^2 \cdot 15}}}$

6 ☐ Ved ikke

Fortsæt på side 20

Opgave XXI

I en pharmaceutisk virksomhed målttes mængden af aktivt stof i 10 tabletter ved 2 forskellige metoder: HPLC og NIR. Resultaterne i gram ses i følgende tabel:

	HPLC	NIR	Difference
Tablet 1	10.4	10.1	0.3
Tablet 2	10.6	10.8	-0.2
Tablet 3	10.2	10.2	0.0
Tablet 4	10.1	9.9	0.2
Tablet 5	11	10.3	0.7
Tablet 6	10.7	10.5	0.2
Tablet 7	10.3	10.2	0.1
Tablet 8	10.9	10.9	0.0
Tablet 9	10.1	10.4	-0.3
Tablet 10	9.9	9.8	0.1

Spørgsmål XXI.1 (29): P-værdien for det tosidede sign test for om der er forskel på målemetodernes niveau bliver:

- 1 ☐ $2 \cdot P(X \geq 6) = 0.2890$, hvor $X \sim b(x; 8, 1/2)$
- 2 ☐ $P(X \geq 6) = 0.1445$, hvor $X \sim b(x; 8, 1/2)$
- 3 ☐ $P(X \geq 8) = 0.0547$, hvor $X \sim b(x; 10, 1/2)$
- 4 ☐ $2 \cdot P(X \geq 8) = 0.1094$, hvor $X \sim b(x; 10, 1/2)$
- 5 ☐ $P(X \geq 6) = 0.3770$, hvor $X \sim b(x; 10, 1/2)$
- 6 ☐ Ved ikke

Opgave XXII

Spørgsmål XXII.1 (30): Hvis man forventer at et tosidet sign test med n observationer får n ens fortegn, hvor mange observationer skal man da planlægge for at man kan forkaste nulhypotesen med $\alpha = 0.01$?

- 1 ☐ 6
- 2 ☐ 7
- 3 ☐ 8
- 4 ☐ $\frac{1}{4} \left(\frac{1.96}{0.01} \right)$
- 5 ☐ $\frac{1}{4} \left(\frac{2.576}{0.01} \right)$
- 6 ☐ Ved ikke