

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Deleksamen i: STK1000 — Innføring i anvendt statistikk.

Eksamensdag: Onsdag 10. oktober 2012.

Tid for eksamen: 15:00 – 17:00.

Oppgavesettet er på 6 sider.

Vedlegg: Svarark.

Tillatte hjelpemidler: Lærebok, ordliste for STK1000, godkjent kalkulator.

Kontroller at oppgavesettet er komplett før du begynner å besvare spørsmålene.

Svarene føres på eget svarark.

Alle 20 oppgaver teller likt. For hver oppgave skal du merke av for bare ett svaralternativ. Du får ett poeng for hvert riktige svar, maksimalt 20 poeng. Dersom du svarer feil eller lar være å krysse av på et oppgave, får du null poeng. Du blir altså ikke "straffet" med minuspoeng for å svare feil.

Lykke til!

Oppgave 1. Vi har følgende datasett med $n = 12$ observasjoner, som er ordnet i stigende rekkefølge

60 80 89 92 98 100 102 103 106 108 121 154

Fem-talls-oppsummeringen for dette datasettet er

A 80 89 100 107 121 **B** 60 90.5 102 108 154

C 60 90.5 101 107 154 **D** 60 89 100 107 154

Oppgave 2. I datasettet i forrige oppgave, hvor mange observasjoner er potensielle uteliggere i i følge $1.5 \times IQR$ -regelen?

A 2 **B** 0 **C** 3 **D** 1

Oppgave 3. Vi har data på CO₂-utslipp per innbygger (variabelen "CO₂") i 26 land i Europa. Under finner du oppsummerende statistikk for datasettet

(Fortsettes på side 2.)

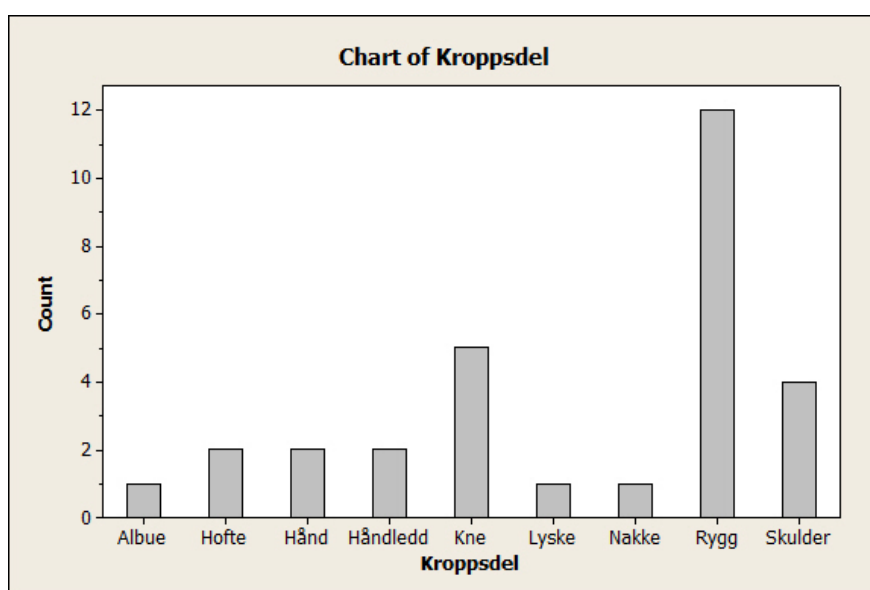
Descriptive Statistics: C02

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
C02	26	0	2,486	0,252	1,283	1,010	1,590	2,165	2,728	6,810

Hva er variansen til observasjonene?

- A** 0.252 **B** 1.283 **C** 0.064 **D** 1.646

Oppgave 4. En fysioterapeut registrerer hvilken kroppsdel hver av 30 tilfeldig utvalgte pasienter kommer for å få behandlet. Under vises resultatet i et stolpediagram.

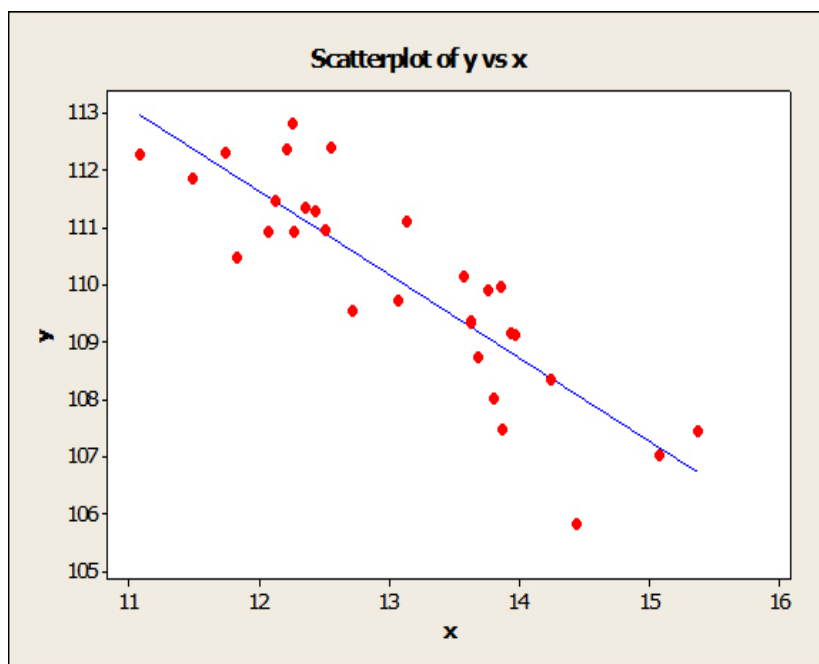


Andelen som kommer for å få behandlet albue, håndledd eller hånd er

- A** 16.7% **B** Mindre enn andelen som kommer til behandling av kne
C Større enn andelen som kommer til behandling av kne **D** 5%

Oppgave 5. Under er et spredningsplott med regresjonslinje for parvise observasjoner av to variabler x og y .

(Fortsettes på side 3.)



Hva er den empiriske korrelasjonen mellom x og y ?

- A -1 B 0.782 C 0.989 D -0.865

Oppgave 6. For å kunne undersøke en eventuell sammenheng mellom oppmøte på forelesninger og eksamensresultater i et statistikkurs ved et universitet, ble poengene på eksamen ("Poeng på eksamen") registrert for 10 studenter med ulike antall (0-9) timers fravær fra forelesning ("Antall fraværstimer"). Under er deler av utskriften av en regresjonsanalyse fra Minitab. Den empiriske korrelasjonen mellom "Antall fraværstimer" og "Poeng på eksamen" er -0.987.

Regression Analysis: Poeng på eksamen versus Antall fraværstimer

Predictor	Coef	SE Coef	T	P
Constant	86,7564	0,8025	108,11	0,000
Antall fraværstimer	-2,5903	0,1503	-17,23	0,000

Hvor mange prosent av variasjonen i "Poeng på eksamen" blir forklart av regresjonen av "Poeng på eksamen" på "Antall fraværstimer"?

- A 99.3% B 98.7% C 0.987% D 97.4%

Oppgave 7. I følge regresjonsanalysen i forrige oppgave, hvor stor endring i predikert antall "Poeng på eksamen" medfører en time økning i "Antall fraværstimer" (avrundet til

(Fortsettes på side 4.)

nærmeste heltall)?

A 1 B -3 C 84 D 87

Oppgave 8. I følge regresjonsanalysen i oppgave 6, hva er predikert antall "Poeng på eksamen" for en student som har 3 timer fravær fra forelesningene (avrundet til nærmeste heltall)?

A 89 B 79 C 95 D 258

Oppgave 9. En forsker skal gjøre et utvalg fra en populasjon av størrelse 100.000, og deretter regne ut en observator. Hvilken utvalgsstørrelse vil gi den minste variabiliteten til observatoren?

A $n = 1$ B $n = 10$ C $n = 100$
D $n = 1$, $n = 10$ og $n = 100$ vil gi samme variabilitet

Oppgave 10. En forsker ønsker å estimere forventet inntekt (μ) til nyutdannede fra UiO. Hun spør 50 tilfeldig utvalgte nyutdannede fra UiO om deres inntekt, og regner deretter ut gjennomsnittet \bar{x} av disse inntektene. Da er

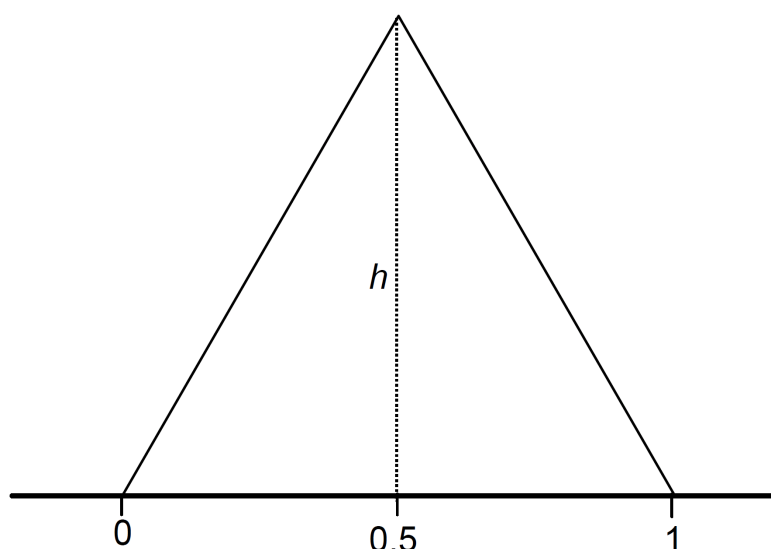
A \bar{x} en parameter B μ en observator C Populasjonsstørrelsen 50
D Utvalgsstørrelsen 50

Oppgave 11. En kontinuerlig tilfeldig variabel X er uniformt fordelt mellom 1 og 10. Hva er $P(X < 5)$?

A 0.50 B 0.40 C 0.44 D 0.56

Oppgave 12. Figuren under viser sannsynlighetstettheten til en kontinuerlig tilfeldig variabel Y .

(Fortsettes på side 5.)



Hva er høyden h i denne sannsynlighetstettheten?

- A** 2 **B** 0.5 **C** 1 **D** 4

Oppgave 13. Utfallsrommet av et eksperiment er $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Vi er interessert i begivenhetene $A = \{2, 4, 7\}$, $B = \{5, 6, 8\}$, $C = \{6, 9, 10\}$, $D = \{2, 3, 4, 7\}$ og $E = \{1\}$. Da er

- A** B og D er uavhengige **B** C og E er uavhengige **C** B og D er disjunkte
D A og D er disjunkte

Oppgave 14. Vi ser fortsatt på utfallsrommet og begivenhetene fra forrige oppgave. Da er

- A** $E^C = A \cup B \cup C$ **B** $E^C = A \cup B \cup D$ **C** $E^C = A \cup C \cup D$
D $E^C = B \cup C \cup D$

Oppgave 15. Vi ser fortsatt på utfallsrommet og begivenhetene fra oppgave 13. Alle utfallene i utfallsrommet S er like sannsynlige. Hva er sannsynligheten for at A eller D skjer?

- A** $\frac{12}{100}$ **B** $P(A)$ **C** $P(D)$ **D** $\frac{7}{10}$

Oppgave 16. Den tilfeldige variabelen X har forventning $\mu_X = 1$ og standardavvik $\sigma_X = 2$. Den tilfeldige variabelen Y har forventning $\mu_Y = 2$ og standardavvik $\sigma_Y = 1$. Hva er forventningen til $Z = X - 2Y$?

(Fortsettes på side 6.)

A -3 B 5 C -1 D -7

Oppgave 17. Vi ser på de samme tilfeldige variablene som i forrige oppgave. Hva er standardavviket til Z ?

A 0 B Har ikke nok informasjon til å kunne regne ut dette C 2.83 D 8

Oppgave 18. Den diskrete tilfeldige variabelen X har forventning 1.5 og sannsynlighetsfordeling

x_i	y	1	4
$P(X = x_i)$	0.2	0.5	0.3

Hva må y være?

A -1 B 0.7 C 0.5 D -3.5

Oppgave 19. Dersom du tipper svaret på hver av de 20 oppgavene i dette eksamenssettet helt tilfeldig, hva er forventet total poengsum?

A 4 B 5 C $\frac{11}{2}$ D 10

Oppgave 20. Dersom du tipper svaret på hver av de 20 oppgavene i dette eksamenssettet helt tilfeldig, hva er standardavviket til total poengsum?

A 3.75 B 0.43 C 8.66 D 1.94

Ferdig!