

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Deleksamen i: STK1000 — Innføring i avvendt statistikk
Eksamensdag: Onsdag 7. oktober 2015
Tid for eksamen: 11.00 – 13.00
Oppgavesettet er på 6 sider.
Vedlegg: Svarark.
Tillatte hjelpemidler: Godkjent kalkulator, lærebok (alle utgaver) og ordliste for STK1000

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

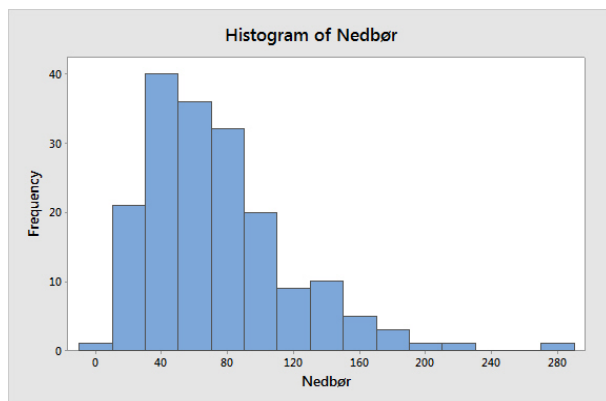
Svarene føres på eget svarark.

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 en oppgave, får du null poeng. Du blir altså ikke straffet med minuspoeng for å svare feil.

Legg merke til at i spørsmålene er noen av svaralternativene avrundet.

Lykke til!

I figuren under finner du et histogram over månedsnedbør i mm på Blindern gjennom årene 2000-2014. På neste side er det satt opp en Minitab-utskrift med deskriptiv statistikk over de samme tallene (men med gjennomsnittet editert bort).



Figur 1: Histogram over månedsnedbør i mm, Oslo, 2000-2014.

(Fortsettes på side 2.)

Descriptive Statistics: Nedbør

Variable	N	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Nedbør	180	3,33	44,68	8,70	40,15	64,45	96,25	279,10

Oppgave 1

Med den formen på fordelingen til nedbørstallene har vi at

- A gjennomsnittet er klart mindre enn medianen
- B gjennomsnittet og medianen er tilnærmet like
- C gjennomsnittet er mindre enn første kvartil
- D gjennomsnittet er større enn medianen

Oppgave 2

For tallene i forrige oppgave er interkvartilavstanden lik

- A 24.30
- B 48.60
- C 56.10
- D 31.80.

Oppgave 3

For nedbørsdataene er variansen lik

- A 11.089
- B 6.68
- C 44.68
- D 1996.302.

Oppgave 4

Hvis vi oppgir nedbørsmengden i cm istedet for mm vil standardavviket bli

- A 4.468
- B 446.8
- C 44.68
- D 3.33.

Oppgave 5

I en studie har man tenkt å samle inn data om totalt 200 personer og det er bestemt at eksakt 100 av disse skal være kvinner. Dette er

- A et stratifisert utvalg
- B et blindet forsøk
- C et randomisert forsøk
- D et enkelt tilfeldig utvalg.

(Fortsettes på side 3.)

Oppgave 6

En diskret tilfeldig variabel X antar verdiene 1, 2 og 3 med sannsynligheter p_1, p_2 og p_3 gitt i følgende tabell

x	1	2	3
sannsynlighet p_x	0.25	0.5	0.25

Da er forventningen μ_X til X lik

- A** 1.5 **B** 2 **C** 2.5 **D** 3.

Oppgave 7

For X gitt i forrige oppgave blir standardavviket lik

- A** 0.25 **B** ca. 0.353 **C** 0.5 **D** ca. 0.707.

Oppgave 8

Anta at temperaturen målt i Celsius en gitt dag er en tilfeldig normalfordelt variabel X med forventning 20 og standardavvik 5, dvs. X er $N(20,5)$. Da er sannsynligheten for at $16 \leq X \leq 24$ lik

- A** ca. 0.21 **B** ca. 0.58 **C** ca. 0.68 **D** ca. 0.95.

Oppgave 9

Anta igjen at temperaturen målt i Celsius en gitt dag X er $N(20,5)$. La $Y = 32 + 1.8X$ være temperaturen målt i Farenheit. Da er Y fordelt som

- A** $N(60,9)$ **B** $N(68,5)$ **C** $N(60,5)$ **D** $N(68,9)$.

Oppgave 10

Anta at temperaturen målt i Celsius to dager er tilfeldige normalfordelte variable X_1 og X_2 begge med forventning 20 og standardavvik 5. Anta videre at X_1 og X_2 er målt med ett års mellomrom slik at de kan regnes som uavhengige. Da er gjennomsnittet $\bar{x} = (X_1 + X_2)/2$ normalfordelt med forventning 20 og standardavvik

- A** 2.5 **B** ca. 3.53 **C** ca. 7.07 **D** 10.

Oppgave 11

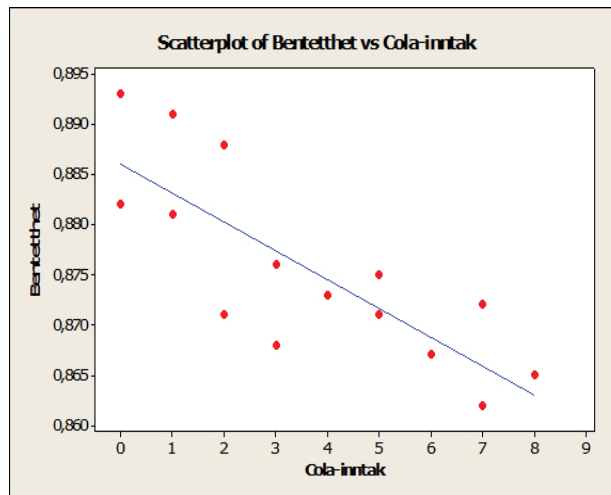
Anta igjen at temperaturen målt i Celsius to dager er tilfeldige normalfordelte variable X_1 og X_2 med forventning 20 og standardavvik 5, men at X_1 og X_2 nå er målt med en dags mellomrom og at korrelasjonen mellom X_1 og X_2 er lik $\rho = 0.62$. Da har gjennomsnittet $\bar{x} = (X_1 + X_2)/2$ forventning 20 og standardavvik

- A** 2.5 **B** ca. 3.53 **C** 4.5 **D** ca. 6.36.

(Fortsettes på side 4.)

Oppgave 12

For å kunne undersøke om inntak av Cola har en sammenheng med bentetthet hos kvinner ble gjennomsnittlig antall Cola-bokser drukket per uke (x) og bentetthet (y) registrert for 15 kvinner. Under ser du et spredningsplott med regresjonslinje for de parvise observasjonene av de to variablene.



Figur 2: Spredningsplott av inntak av Cola mot bentetthet hos kvinner

Den empiriske korrelasjonen mellom Cola-inntak og bentetthet er ett av tallene oppgitt under. Angi hvilket.

- A** 0.92 **B** -0.81 **C** -1.00 **D** 0.53

Oppgave 13

Under er deler av utskriften av en regresjonsanalyse fra Minitab for datasettet fra forrige oppgave

Regression Analysis: Bentetthet versus Cola-inntak

The regression equation is

Bentetthet = 0,886 - 0,00289 Cola-inntak

Predictor	Coef	SE Coef	T	P
Constant	0,886068	0,002600	340,82	0,000
Cola-inntak	-0,0028893	0,0005892	-4,90	0,000

En kvinne drikker 4 bokser Cola i uken. Hva er predikert bentetthet for denne kvinnen?

- A** 0.8744 **B** 0.8861 **C** 0.8976 **D** 0.00289

(Fortsettes på side 5.)

Oppgave 14

I følge regresjonsanalysen i forrige oppgave, hvor stor endring i predikert benteitet medfører en økning i Cola-inntak på en boks i uken?

- A** -0.00289 **B** 0.8861 **C** 0.00289 **D** -0.8861

Oppgave 15

I en Non Stop-pose er det brune, grønne, gule, røde og svarte Non Stop. Du skal ta en Non Stop ut av posen. La A være at den er gul eller grønn. La B være at den ikke er rød. La C være at den er svart, brun eller rød. Da er

- A** A og B disjunkte
B A og C disjunkte
C B og C disjunkte
D B og C^c disjunkte

Oppgave 16

Vi ser på de samme hendelsene som i forrige oppgave. Der er

- A** $A^c = C$
B $B^c = A^c$ og C
C $B^c = A^c$ eller C
D $B^c = A$ og C

Oppgave 17

Du får oppgitt at $P(A) = 0.59$, $P(B) = 0.27$ og $P(B | A) = 0.45$. Hva er $P(A \text{ og } B)$?

- A** 0.1593 **B** 1.04 **C** 0.86 **D** 0.2655

Oppgave 18

I 2005 var 12.64% av fødslene i USA premature (for tidlig fødte). I den samme perioden var det 0.22% av fødslene der babyen veide 4 kg eller mer og som samtidig var premature. Hva er sannsynligheten for at en baby veier 4 kg eller mer ved fødsel, gitt at den blir født prematurt?

- A** 0.00028 **B** 0.57 **C** 0.1286 **D** 0.0174

(Fortsettes på side 6.)

Oppgave 19

Med uavhengige og identiske fordelte tilfeldige variable X_1, \dots, X_n med forventning μ vil gjennomsnittet \bar{x} gå mot μ når n vokser. Dette resultatet kalles

- A sentralgrenseteoremet
- B et randomisert eksperiment
- C Bayes regel
- D store talls lov.

Oppgave 20

Med uavhengige og identiske fordelte tilfeldige variable X_1, \dots, X_n med forventning μ og standardavvik σ vil gjennomsnittet \bar{x} være tilnærmet normalfordelt når antall observasjoner n er stort. Dette resultatet kalles

- A sentralgrenseteoremet
- B et randomisert eksperiment
- C Bayes regel
- D store talls lov.

SLUTT