

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i: STK 1000 — Innføring i anvendt statistikk.

Eksamensdag: Mandag 30. mai 2005.

Tid for eksamen: 09.00 – 12.00.

Oppgavesettet er på 4 sider.

Vedlegg: Tabell over normalfordeling, tabel over t -fordeling.

Tillatte hjelpemidler: Lærebok: Moore & McCabe "Introduction to the practice of statistics", 3. el. 4. utgave. Ordliste for bruk i STK 1000, kalkulator.

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

Oppgave 1.

Utslippmengden av karbondoksyd (CO) fra biler av et bestemt bilmerke kan betraktes som tilfeldige variable som er uavhengige og normalfordelte med samme forventning og standardavvik 2 gram/mil.

- a) Produsenten hevder at sannsynligheten for at utslippet overstiger 16 gram/mil for en bil, høyst er 0.05. Hva må forventningen da være?
- b) En bilforhandler selger 10 biler av denne typen i løpet av en måned. Hva er fordelingen til antallet biler som slipper ut mer enn 16 gram/mil? Begrunn svaret.

(Fortsettes side 2.)

Oppgave 2.

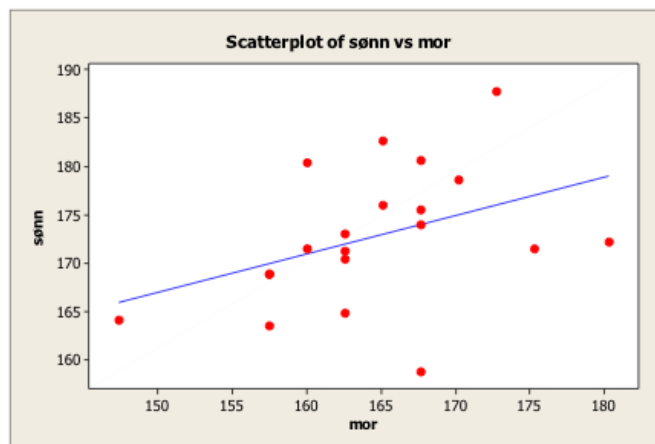
Målinger av antall kalorier pr. gram i 16 tilfeldig utvalgte frokostblandinger gav følgende resultat for gjennomsnitt, \bar{x} , og empirisk standardavvik, s .

Antall	Gjennomsnitt	Standardavvik
16	3.7625	0.2217

- Finne et 95% konfidensintervall for forventet antall kalorier. Hva er tolkningen av et slikt intervall?
- Forklar hvilke signifikanstester som kan utføres ved hjelp av konfidensintervallet du fant i punkt a).
- Det er av interesse å finne ut om forventet antall kalorier i slike frokostblandinger er større enn 3.6 pr. gram. Formuler en nullhypotese og alternativ hypotese for å besvare dette. Beregn en øvre og nedre grense for P-verdien. Hva er konklusjonen på testen?

Oppgave 3.

Spredningsplottet nedenfor viser høyden i cm for 20 menn og deres mødre.



MINITAB-utskriften nedenfor viser resultatet av å tilpasse en regresjonslinje ved minste kvadraters metode.

(Fortsettes side 3.)

The regression equation is
 sønn = 108 + 0.394 mor

Predictor	Coef	SE Coef	T	P
Constant	107.90	34.55	3.12	0.006
mor	0.3941	0.2099	1.88	0.077

S = 6.65277 R-Sq = 16.4% R-Sq(adj) = 11.7%

Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	174.89	1.89	(170.92, 178.85)	(160.36, 189.42)

Values of Predictors for New Observations

New Obs	mor
1	170

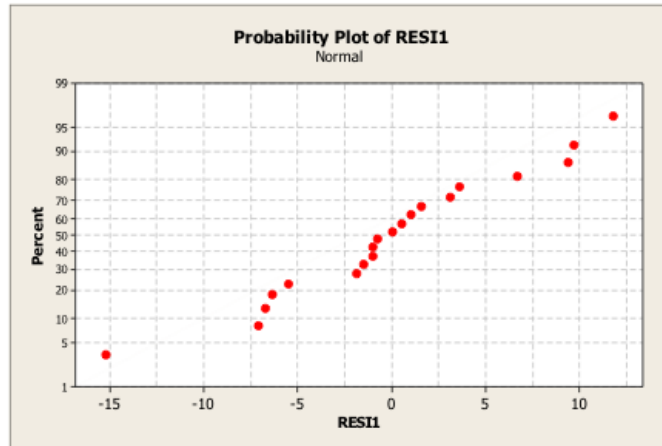
- Formuler en regresjonsmodell for dette problemet. Presiser hva de enkelte variablene betyr og hva parametrene er. Bruk utskriften til å finne estimater for parametrene.
- La β_1 betegne stigningsforholdet for regresjonslinja i populasjonen. Test nullhypotesen mot den alternative ensidige hypotesen

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 > 0.$$

Forklar hvordan P-verdien beregnes. Hva er den i dette tilfellet, og hva er konklusjonen?

- Lag et 90% konfidensintervall for stigningsforholdet, β_1 .
- Utskriften viser også et 95% konfidensintervall for den forventede verdien og et prediksjonsintervall for høyden til sønner av mødre som er 170 cm høye. Forklar hva tolkningen av disse intervallene er. Hvorfor er prediksjonsintervallet lengst?
- Figuren nedenfor viser et normalfordelingsplott for residualene. Forklar hvordan plottet beregnes og hva det brukes til. Hva er konklusjonen i dette tilfellet? Hva er omtrentlig høyde for sønnen og moren svarende til residuallet lengst til venstre i figuren?



SLUTT