

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i STK1005 — Innføring i anvendt statistikk B

Eksamensdag: 16. juni 2009

Tid for eksamen: 9.00–13.00

Oppgavesettet er på 4 sider.

Vedlegg: Kap. 15 fra bokens CD-rom.

Tillatte hjelpemidler: Læreboken Moore & McCabe: Introduction to the practice of statistics, godkjent kalkulator

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

Oppgave 1

En artikkel i 'Journal of Abnormal Psychology' rapporterer bl.a. følgende forsøk for å undersøke effekten av hypnose som smertestiller: $n = 8$ pasienter oppgir smertenivå ved varmebelastning på hud før og etter hypnose, der høy verdi betyr høyt smertenivå (målt i cm. på en smerteskala). Målingene var

| Person | Before | After | Diff |
|--------|--------|-------|------|
| 1 | 6.6 | 6.8 | -0.2 |
| 2 | 6.5 | 2.4 | 4.1 |
| 3 | 9.0 | 7.4 | 1.6 |
| 4 | 10.3 | 8.5 | 1.8 |
| 5 | 11.3 | 8.1 | 3.2 |
| 6 | 8.1 | 6.1 | 2.0 |
| 7 | 6.3 | 3.4 | 2.9 |
| 8 | 11.6 | 2.0 | 9.6 |

og artikkelen konkluderer med at smertemålingene er signifikant lavere etter hypnose.

a) Det er naturlig å parre observasjonene, dvs. se på differansene før - etter (Before - After) for hver person for å si noe om endring i smertenivå ved hypnose. Anta at disse differansene kommer fra en normalfordeling. Finn estimer for forventningen μ og standardavviket σ i denne fordelingen. Du kan hente tall fra Minitabutskriften nedenfor.

b) Konstruer et 95% konfidensintervall for μ . Forklar kort forskjellen på et 95% og 99% konfidensintervall.

(Fortsettes på side 2.)

Paired T for Before - After

| | N | Mean | StDev | SE Mean |
|------------|---|-------|-------|---------|
| Before | 8 | 8.712 | 2.177 | 0.770 |
| After | 8 | 5.588 | 2.608 | 0.922 |
| Difference | 8 | 3.13 | 2.91 | 1.03 |

c) Forfatterne har testet hypotesene $H_0 : \mu = 0$ mot $H_a : \mu > 0$. Bruk signifikansnivå 0.05 og vis at det er riktig at smertenivået er signifikant lavere etter hypnosen, dvs. at vi kan forkaste H_0 og konkludere med H_a .

d) Betyr dette at man har bevist at hypnose har smertestillende effekt? Kritiser forsøksopplegget med vekt på placeboeffekt.

e) Test til slutt om smertenivået er signifikant lavere etter hypnose ved hjelp av en passende ikkeparametrisk test. Spesifiser hvilken test du bruker, og vis hvordan du har regnet ut P-verdien.

Oppgave 2

I et randomisert, dobbelt blindt, sammenlignende forsøk fikk 1070 barn en influensavaksine i form av nesenspray, og 532 barn placebo (også i form av nesenspray). 128 av barna som fikk vaksine fikk senere influensa, mot 95 i placebogruppen. Du skal finne ut om disse tallene gir belegg for å påstå at influensavaksinen har positiv effekt.

a) Først skal du svare på dette ved å sammenligne andelene som får influensa i de to populasjonene (vaksine og placebo). Kall disse p_1 og p_2 . Sett opp passende nullhypotese og alternativ hypotese og formel for tilhørende z-observator. Beregn numerisk verdi for denne.

b) Finn P-verdi for testen og konkluder. Kommenter gyldigheten av tilnærmingen du har brukt.

c) Sett deretter opp datasettet i en toveistabell og utfør en χ^2 -test for å finne ut om det er noen sammenheng mellom vaksinerings og forekomst av influensa.

d) Sammenlign resultatene du fikk i b) og c) og forklar sammenhengen.

(Fortsettes på side 3.)

Oppgave 3

Forskjellige frokostblandinger inneholder ulik sammensetning av fett og sukker. Mengden fett og sukker (begge målt i gram per gram frokostblanding) sammenholdes med kalori-innholdet i blandingen (målt i kalorier per 100 gram frokostblanding) for 16 forskjellige frokostblandinger til salgs i vanlige forretninger. Formålet med analysen er å forstå i hvilken grad kalori-innholdet bestemmes av og kan predikeres fra fett- og sukker-innholdet i blandingene.

- Sett opp modell og antakelser for en multippel lineær regresjon med kalori-innhold som respons og fett- og sukker-innhold som forklaringsvariable. Finn estimater for alle parametre i modellen fra Minitabutskriften.
- Sett opp hypotesene som testes og vis hvordan t-observator og P-verdi for regresjonskoeffisienten til forklaringsvariabelen fett (Fat) er beregnet. Spesifiser spesielt antall frihetsgrader i t-fordelingen du bruker.
- Bruk utskrift og plott til å kommentere hvor godt modellen passer til data og i hvilken grad den egner seg til prediksjon.
- Finn et estimat for forventet kalori-innhold i en blanding med 0.05 g fett og 0.25 g sukker per gram frokostblanding. Finn et 95% konfidensintervall for samme størrelse.
- Minitab har også beregnet et 95 % prediksjonsintervall for kombinasjonen av fett og sukker i d). Forklar hva dette prediksjonsintervallet representerer, og skisser hvordan det er beregnet.

Regression Analysis: Calories versus Fat, Sugar

The regression equation is
 Calories = 3.50 + 3.86 Fat + 0.809 Sugar

| Predictor | Coef | SE Coef | T | P |
|-----------|---------|---------|--------|-------|
| Constant | 3.49895 | 0.03333 | 104.98 | 0.000 |
| Fat | 3.8619 | 0.8734 | 4.42 | 0.001 |
| Sugar | 0.8087 | 0.1002 | 8.07 | 0.000 |

S = 0.0605364 R-Sq = 89.6% R-Sq(adj) = 88.0%

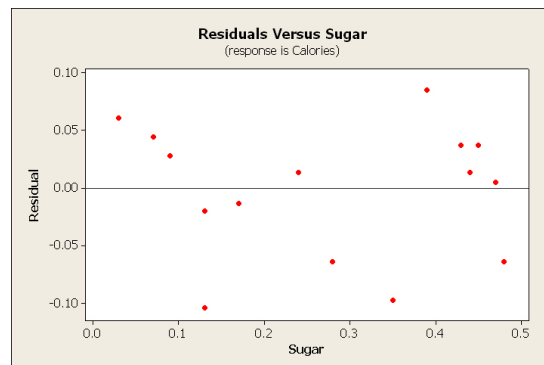
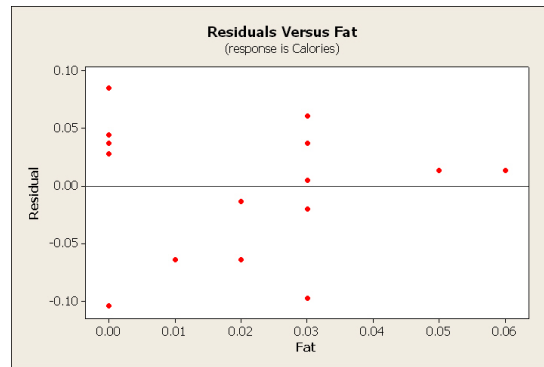
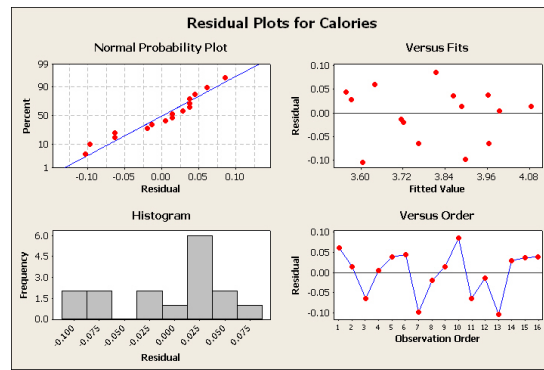
Predicted Values for New Observations

| New Obs | Fit | SE Fit | 95% CI | 95% PI |
|---------|--------|--------|------------|------------------|
| 1 | 3.8942 | 0.0304 | (???, ???) | (3.7479, 4.0405) |

Values of Predictors for New Observations

| New Obs | Fat | Sugar |
|---------|--------|-------|
| 1 | 0.0500 | 0.250 |

(Fortsettes på side 4.)



SLUTT