UNIVERSITETET I OSLO

DET MEDISINSKE FAKULTET

Eksamen

ERN4000 STATISTIKK

Fredag 6. september 2024 kl 09:00-12:00

Hjelpemidler

- Stata (på den stasjonære maskinen)
- Egen kalkulator (med exp og ln)
- Stata Cheat Sheet
- Lærerbok: Aalen OO (red), Frigessi A, Moger TA, Scheel I, Skovlund E, Veierød MB. Statistiske metoder i medisin og helsefag. Gyldendal Akademisk 2006.
  eller
  Aalen OO (red), Frigessi A, Moger TA, Scheel I, Skovlund E, Veierød MB. Statistiske metoder i medisin og helsefag, 2. utgave. Gyldendal Akademisk 2018.

Kort forklaring til oppgavene

I oppgavene skal du foreta statistisk analyse av noen medisinske problemstillinger. I presentasjonen av resultatene er det tre statistiske mål som er viktige å få frem: estimat (f.eks. gjennomsnitt, andel, odds ratio, regresjonskoeffisient), konfidensintervall og p-verdi.

Datafilen nhanes.dta brukes i oppgave 1, 2 og 3

Datafilen er fra data samlet inn av US National Center for Health Statistics. Data fra 750 deltakere er inkludert med følgende variabler:

Variabelnavn	**Beskrivelse**
Age	Alder (år)
Gender	Kjønn (1=kvinne, 2=mann)
Weight	Vekt (kg)
Height	Høyde (cm)
TotChol	Kolesterol (mmol/L)
Diabetes	Diabetes (1=nei, 2=ja)
PhysActive	Fysisk aktivitet (1=nei, 2=ja)

Oppgave 1

a) Gi en beskrivelse av alder, vekt, høyde og kolesterol. Bruk det du synes er mest hensiktsmessig av beskrivende statistikk og presenter resultatene i en tabell.
b) Lag tre tabeller: en for kjønn, en for diabetes og en for fysisk aktivitet. Hver tabell skal vise antall personer og prosentandel i hver av variablenes kategorier, samt totalsummen.
c) Sammenlign høyden for menn og kvinner ved hjelp av en statistisk test. Begrunn valget av test. Kommenter kort det du finner.
d) Lag et spredningsdiagram for sammenhengen mellom høyde (x-akse) og vekt (y-akse), og beregn korrelasjonskoeffisienten for høyde og vekt. Begrunn valget av korrelasjonskoeffisient. Kommenter kort det du finner.

Oppgave 2

Vi fortsetter med datafilen nhanes.dta, og ser på sammenhengen mellom fysisk aktivitet og diabetes. Fysisk aktivitet er eksponering og diabetes er sykdom.

a) Lag en tabell med diabetes i radene og fysisk aktivitet i kolonnene. I tillegg til å inkludere antall i hver celle skal tabellen prosentueres innen hver kolonne. Kommenter kort det du finner.
b) Vi fortsetter med tabellen i punkt a). Beregn risikodifferansen med 95% konfidensintervall og relativ risiko (risk ratio) med 95% konfidensintervall. Kommenter kort det du finner.
c) Hvilken test vil du bruke for å undersøke om det er signifikant sammenheng mellom fysisk aktivitet og diabetes? Begrunn svaret. Utfør testen og oppgi p-verdien fra testen. Kommenter kort det du finner.

Oppgave 3

Vi fortsetter med datafilen nhanes.dta.

a) Forskerne ville undersøke om kolesterol nivået er forskjellig hos menn og kvinner. Hvilken test vil du bruke for å undersøke dette? Utfør testen og kommenter kort det du finner.
b) Forskerne var interessert i om sammenhengen mellom kjønn og kolesterol ble annerledes hvis de utførte en multippel regresjonsanalyse der de justerte for alder og vekt. Utfør en slik multippel regresjonsanalyse (kolesterol er 'independent variable'). Er kjønn signifikant i denne modellen? Oppgi regresjonskoeffisienten for kjønn med 95% konfidensintervall. Hvordan tolkes den estimerte regresjonskoeffisienten for kjønn?
c) Hva er R^2 for den multiple regresjonsmodellen? Hvordan tolkes denne?
d) Var forutsetningene til stede for å gjøre analysene i b)? Begrunn svaret.

Oppgave 4

Multippel sklerose (MS) er en potensielt invalidiserende sykdom. Noen pasienter med alvorlig MS kan miste evnen til å gå, mens andre kan oppleve milde symptomer som nummenhet eller svakhet i lemmer, tretthet eller kognitive problemer.

Forskjellige diagnostiske metoder, inkludert magnetresonanstomografi (MR), brukes til å diagnostisere sykdommen. For å bekrefte sykdommen er det imidlertid nødvendig med detaljert sykehistorie, nevrologisk undersøkelse, blodprøver og cerebrospinalvæskeanalyse.

I en hypotetisk klinikk ble det utført MR for 50 pasienter. Til slutt ble 32 pasienter bekreftet å ha MS, og av disse hadde 25 positive MR-resultater. Blant de som ikke har MS, hadde 3 pasienter positive MR-resultater.

Fyll inn antallene i tabellen nedenfor, og beregn sensitivitet og spesifisitet for MR.

	MS		
MR resultat	Ja	Nei	Totalt
Positiv	25	3	
Negativ			
Totalt	32		50