**Stat100 vår 2018**

**Oppgaver til uke 14**

**Data til oppgave ligger som Exel-fil. Last denne ned og bruk klipp og lim.**

**Oppgave 1A og Oppgave 1B. Besvares med hypotesetesting og P-verdi.**

**Sett opp null hypotese og alternativ hypotese. Bruk tabell for binomisk sannsynlighet.**

**Husk: P-verdi er P(resultat eller bedre i favør av H1 hvis H0 er sann).**

**Hva er Type I og Type II feil , vær konkret.**

**Oppgave 1A**

A får 100 kroner for hver mynt, B får 100 for hver kron. Vi kaster 10 ganger med en mynt som B kommer med. Det blir 9 kron. Vil du anmelde B for å ha en juksemynt?

**Oppgave 1B**

Dersom du innfører en ny plantesort og du får 9 av 10 som spirer, kan du påstå med rimelig sikkerhet at denne nye sorten er bedre enn den nåværende som har spireprosent 70?

**Oppgave 2** (Bruk tilnærming til normalfordeling).

Du vurderer å importere en ny plantesort, men bare dersom du er rimelig sikker på at den har bedre spireprosent enn 80.

Dersom 90 av 100 planter spirer, vil du importere plantesorten? Svar med P-verdi.

**Oppgave 3 (Lur oppgave Om t-fordeling. Bruk tabellen i boka)**

Anta at du tar 4 uavhengige observasjoner fra en normalfordeling, der begge parametre (både  og ) er **ukjente.** Disse 4 observasjonene ga et utvalgsgjennomsnitt på 3 og et utvalgs-standardavvik på 1.

Lag et 90 % og 95 % KI for forventningen.

Hvis du vil teste en hypotese om at forventning er 2 mot et ensidig alternativ på at den er større enn 2, på hvilket signifikansnivå kan du forkaste hypotesen?

**Oppgave 4*Hastighet på biler***

På en vegstrekning har det i lang tid vært en fartsgrense på 50km/t. Fartsovervåkning av strekningen over svært lang tid indikerer at bilers hastighet er normalfordelt med forventning 52km/t.

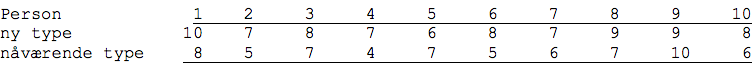
Hastigheten på strekningen ble så satt opp til 60 km/t. Noen dager etter fartshevningen ble hastigheten til 5 tilfeldig valgte biler målt til:

63, 58, 45, 70, 66

Du får i oppgave å teste om hevingen fra 50 til 60 km/t har ført til en forventet økning i fart på strekningen.

**Oppgave 5 (tidligere eksamensoppgave, men lurt å bruke R-commander)**

Et firma ønsker å teste ut enn ny hamburgertype. De trekker ut 10 personer tilfeldig og lar hver av disse smake på den nye hamburgeren, for deretter å gi poeng på en skala fra 0 til 10. For å ha noe å sammenligne med lar de også de samme personene smake på en hamburger av den nåværende typen og gi poeng til denne. (10 poeng er best).



**a)** Det ble utført en statistisk hypotesetest (t-test) der alternativet var om folk likte den nye typen best.

Sett opp en modell for data.

Sett opp nullhypotese og alternativ.

Finn p-verdien til denne testen.

Forklar hamburgerprodusenten, som ikke har kunnskap i statistikk, hva denne P-verdien betyr i praksis.

**b)   (løses uten R-commander)**

En annen måte å teste den nye hamburgertypen på, er å telle hvor mange som sier den nye typen er best. La X være antallet av de 10 som likte ny type best.

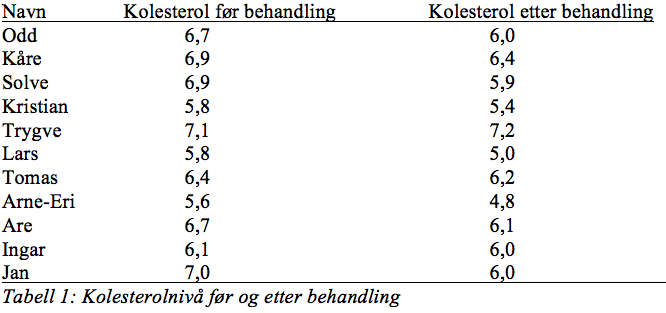
Hva slags sannsynlighetsfordeling har X?

Sett opp en statistisk hypotesetest basert på X, der alternativet er om folk foretrekker den nye typen.

Finn p-verdi på denne testen. Hvorfor tror du at p-verdien blir større i denne testen sammenlignet med den i oppgave a?

**Oppgave 6 (kan løses uten R.commander):**

Høyt kolesterolnivå er et stort problem, dette gjelder særlig for middelaldrende menn med en gjennomført usunn livsstil. Personene under er trukket ut fra denne (særdeles nitriste) populasjonen, og de fikk målt kolesterolinnholdet i blodet (måleenhet er mmol/l). Som et forsøk ble etter loddtrekning Odd, Kristian, Trygve, Lars, Tomas og Jan satt på en streng kostholdsdiett som ernæringsspesialister mente ville redusere kolesterolnivået, mens de resterende ble satt på et program med mye mosjon. Ett år etter at tiltakene ble satt i sving ble kolesterolnivået målt på nytt:



Med positiv effekt mener vi at det vil redusere gjennomsnittlig kolesterolnivå i hele populasjonen.

1. Kan du på bakgrunn av dette påstå at rett kosthold har positiv effekt?

Sett opp modellen du bruker.

Svar ved hypotesetester med signifikansnivå 5 %.

1. Kan du på bakgrunn av dette påstå at mye mosjon har positiv effekt?

Sett opp modellen du bruker.

Svar ved hypotesetester med signifikansnivå 5 %.

1. Kan du påstå at det er signifikant forskjell mellom kosthold og mosjon med hensyn på kolesterolreduksjon?

**Oppgave 7 (tidligere eksamensoppgave)**

Medisinen Nembuxil påstås å virke mot migrene hos kvinner som er plaget av dette. Anta at for denne populasjonen kvinner er antallet migreneanfall i løpet av et år tilnærmet normal-fordelt. For å teste ut medisinen ble 9 kvinner som alle hadde vært plaget av migrene trukket ut. Disse fikk registrert antall migreneanfall i ett år. Dette ga følgende:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Kvinne | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Antall | 12 | 21 | 21 | 9 | 14 | 6 | 20 | 18 | 17 |

*Tabell 1a. Antall migreneanfall i året før kvinnene fikk medisin.*

Året etter gikk de samme 9 på Nembuxil og fikk registrert antall migreneanfall i dette året. Dette ga følgende:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Kvinne | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Antall | 5 | 17 | 18 | 7 | 14 | 5 | 11 | 17 | 18 |

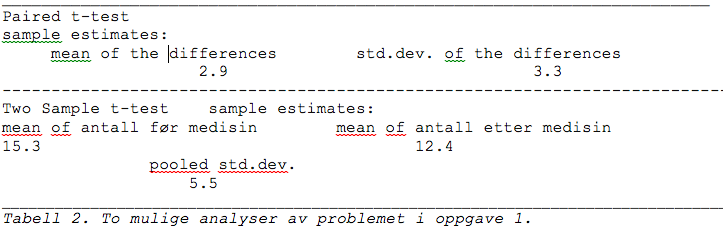
*Tabell 1b. Antall migreneanfall i året der kvinnene fikk medisin.*

Det er vedlagt to mulige analyser for dette forsøket (se Tabell 2).

**a)**   Hvilken av de to analysene er den eneste akseptable å bruke? Svaret må begrunnes.

Estimer forventet effekt av Nembuxil med hensyn på reduksjon av migreneanfall.

**b)**   Har vi bevist at Nembuxil har positiv effekt? Svar ved hypotesetesting. Velg signifikansnivå selv. Finn p-verdien til testen (bruk R-commander)



**Oppgave 8.**

En psykolog påstår at det er systematiske forskjeller i aggressivitet mellom eneggede tvillinger avhengig av hvem som ble født først. Åtte tvillingpar ble gitt en psykologisk test som måler aggressivitet, der høy verdi betyr stor aggressivitet. Dette gav følgende resultat:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tvillingpar nr. 1 | Tvillingpar nr. 2 | Tvillingpar nr. 3 | Tvillingpar nr. 4 | Tvillingpar nr. 5 |
| Født først: 88 | Født først: 74 | Født først: 82 | Født først: 76 | Født først: 93 |
| Født sist: 82 | Født sist: 72 | Født sist: 82 | Født sist: 77 | Født sist: 92 |

|  |  |  |
| --- | --- | --- |
| Tvillingpar nr 6 | Tvillingpar nr 7 | Tvillingpar nr 8 |
| Født først: 66 | Født først: 85 | Født først: 90 |
| Født sist: 70 | Født sist: 81 | Født sist: 86 |

Finn et 95% konfidensintervall for forventet forskjell i aggressivitet og bruk intervallet til å teste psykologens påstand.

# Oppgave 9

**Litt fra *Flervalgsdelen av eksamen* vår 2016**

For Oppgavene **F1 - F4**

Dette er basert på artikkelen: *The antihypertensive effects of fish Oil; av Knapp & FitzGerald publisert i New England Journal of Medicine 1989.*

14 menn med høyt blodtrykk ble tilfeldig valgt ut, og deretter basert på loddtrekning ble 7 av disse satt på en diett basert på fiskeolje, mens de resterende 7 gikk på en diett som brukte helt vanlig matolje. Reduksjonen i diastolisk blodtrykk (målt i Torr, en målestokk for trykk) sammen med deler av en R-commander utskrift er gitt i Tabell 5.

La Xi være blodtrykksreduksjon for person nummer i på fiskeolje og Yi blodtrykksreduksjon for person nummer i på matolje. Anta modellen Xi ~ N(*1*,**), og Yi ~ N (*2*,**).

Mann nr. 1 2 3 4 5 6 7

Diett med fiskeolje: 8 12 10 14 2 0 0

Diett med vanlig olje: -6 0 1 2 -3 -4 2

mean sd n

Fiskeolje 6.57 5.85 7

Vanlig olje -1.14 3.18 7

t = 3.0621, df = 12, p-value = 0.005

alternative hypothesis: true difference in means greater than 0

*Tabell 5: Reduksjon i blodtrykk samt deler av R-commander utskrift*

**F1**

Estimert forventet gevinst (i form av redusert blodtrykk) ved bruk av fiskeolje framfor vanlig olje er:

A) 5,43 B)-5,43 C) 6,57 D) 2,03 E) 0 F) 7,71

**F2**

Hva er det beste estimatet for**?

A) 4,51 B) 4,71 C) 5.85 D) 22,18 E) 2,67 F) 7,6

**F3**

Testing av

H0: Valg av oljetype har ingen betydning for blodtrykksreduksjon, mot

H1: Fiskeolje har positiv betydning for blodtrykksreduksjon

kan formuleres som:

A) H0: *1* = *2* H1: *1* ≠ *2*

B)H0: *1* = *2* H1: *1* > *2*

C) H0: *1* = *2* H1: *1* < *2*

D) H0: *1* ≠ *2* H1: *1* = *2*

E) =  H1:  > 

F)  >  H1:  > 

**F4**

Hva er rett?

A) Hvis *1* = *2*, så er P(- ≥ 7,71) = 0,005

B) Hvis -≥ 7,71, så er P(*1* = *2*,) = 0,005

C) Hvis *1* > *2*, så er P(-) = 0,005

D) P(*1* = *2*) = 0,005

E) P(*1* > *2*) = 0,995

F) Hvis= , så er P(*1* = *2*) = 0,005

**I tillegg.**

Tolk alle tre parametere i modellen.

Hvilken konklusjon trekker du?