**Stat100 vår 2018**

**Løsningsforslag til uke 14**

**Oppgave 1A**

La X være antall kron i 10 kast. Da er X binomisk fordelt med n = 10 og p = P(kron i ett kast).

H0: p = 0,5 (jukser ikke)

H1: p > 0,5 (jukser)

P-verdi: P(X ≥ 9| H0 er sann ).

Dette er 1  P(X ≤ 8) = 1 0,989 = 0,011 (Tabell i boka).

Kan forkaste H0 på 5 % nivå. Påstår juks.

Type I feil: Påstå juks, dersom det ikke er juks.

Type II feil: La være å påstå juks til tross for at det jukses.

--------------------------------------------------------------------------------------------------------

**Oppgave1B**

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?

La X være antall av 10 nye som spirer. Da er X binomisk fordelt med n = 10 og

p = P(en plante spirer).

H0: p = 0,7

H1: p > 0,7

P-verdi: P(X ≥ 9| H0 er sann ).

Dette er 1  P(X ≤ 8) = 1  0,851 = 0,149

Kan ikke forkaste H0 på 5 % nivå (eller på noe akseptabelt nivå).

Har ikke klart å påvise at ny sort er bedre (har høyere spireprosent) enn den gamle.

Type I feil: Innføre ny plantesort som ikke har bedre spireevne en den nåværende.

Type II feil: La være å innføre ny sort med bedre spireevne enn den nåværende.

**Oppgave 2**

La X være antall av 100 nye som spirer.

Da er X binomisk fordelt med n = 100, og p = P(en plante spirer).

H0: p = 0,8

H1: p > 0,8

Husk at når n er stor så vil ~ N(),

Vi får **** = 0,9

Anta at H0 er sann. Da har vi: P(**> 0,9**) = P(**)** =

P(Z > 2,5) = 0,006.

Kom igjen! Start importen av ny plantesort. Vi er temmelig sikre på at denne er bedre.

**Oppgave 3**

Vi har 3 frihetsgrader (4 - 1), da er t0,05;3 = 2,353, og et 90 % KI er

3 ± 2,353\*1**/** = **(1,823; 4,177)**

Vi har 3 frihetsgrader (4-1) da er t 0,025;3 = 3,182, og et 95 % KI er

3±3,182\*1**/ = (1,409; 4,591)**

H0:  = 2 H1:  > 2

T = ** = 2**

Vi har: t 0,05; 3 = 2,353, og t0,1;3 = 1,638.

Vi kan forkaste på 10 %, men ikke på 5 % nivå. P-verdien ligger mellom 0,05 og 0,10.

**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.

La Xi være hastighet til bil nr. i (og dette etter omlegging).

Antar modellen: Xi ~N(, ) og at alle Xi er uavhengige.

Tester:

H0:  = 52 (ingen endring) mot H0:  > 52 (fartsøkning)

Finner T =  = 1,94. Dette må sjekkes mot tabellverdi for t-fordeling med 4 frihetsgrader. Kan forkaste på 10 % , men ikke på 5 % nivå. Altså må p-verdien ligge mellom 0,05 og 0,1.

R-comander

One Sample t-test

t = 1.9446, df = 4, p-value = 0.06186

alternative hypothesis: true mean is greater than 52

sample estimates:

mean of x std.dev. of x

60.400000 9.659193

**Oppgave 5 (R-commander)**

**a)** Modell La Xi være poeng for ny type av person i og Yi være poeng for nåværende type fra person i. (samme person).

Di = Xi - Yi. Vi antar at Di ~N(D, D) og at alle Di er uavhengige.

H0: D = 0 mot H0: D > 0

Paired t-test

data: hamburger$ny and hamburger$gammel

t = 3.0963, df = 9, **p-value = 0.0064**

alternative hypothesis: true difference in means is greater than 0

sample estimates:

mean of the differences std.dev. of the differences

1.400000 1.429841

**P-verdi:** Dersom befolkningen ikke synes det var smasksforskjeller mellom hamburgertypene er det bare 6,4 promilles sjanse for at personene i utvalget skulle ha så sterk (eller sterkere) preferanse for den nye typen.

**b)   (løses uten R-commander)**

X er binomisk fordelt med n= 10 og p = P(en tilfeldig valgt person synes ny type er best).

X er 8, men p er ukjent, men estimers til 0,8

H0: p = 0,5 (ingen forskjell mellom hamburgertypene)

H1: p > 0,5 (flertall i befolkningen fortrekker ny type)

P-verdi: P(X ≥ 8) hvis H0 er sann.

Dette er 1 - P(X ≤ 7) = 0,055.

Kan ikke forkaste H0 på 5 % nivå, men på 6 % nivå.

Tar ikke hensyn til hvor sterk preferansen er, kun om de liker den nye beder enn den gamle.

**Oppgave 6:**

1. Parvis sammenligning fordi hver person utgjør et par.

**Kosthold:** La Xi, være kolesterol for person i før diettstart, og Yi være kolesterol året etter for person i, der i = 1, 2, 3, 4, 5, 6.

Hvis både X og Y er normalfordelte vil: Di = Xi - Yi, også være normalfordelt: Dermed vil

D1, .  . . ., D6 være uavhengig og normalfordelt med (\mu_d, \sigma_d),  der \mu_d = \mu_x - \mu_y.

Vi antar at \sigma_d er ukjent.

Reduksjon i kolesterol ved kosthold

Odd                 0,7

Kristian            0,4

Trygve            -0,1

Lars                 0,8

Tomas             0,2

Jan                 1,0

H0: Ingen effekt av kosthold               \mu_d = 0

H1: positiv effekt av kosthold             \mu_d > 0

T=\frac{\bar{D}}{S/\sqrt{n}}=\frac{0.5}{0.41/\sqrt{6}}=2.99.

Denne er t-fordelt med 5 frihetsgrader under H0.

Dette må sammenlignes med en passende tabellverdi. På 5 % nivå er dette 2,015. Altså kan vi forkaste H0, og påstå at rett kosthold reduserer kolesterolnivået.

1. Mosjon: Modell, etc. som for kosthold:

Reduksjon i kolesterol ved mosjon:

Kåre                 0,5

Solve                1,0

Arne-Eri          0,8

Are                  0,6

Ingar                0,1

H0: Ingen effekt av mosjon                 \mu_d = 0

H1: positiv effekt av mosjon               \mu_d > 0

T=\frac{\bar{D}}{S/\sqrt{n}}=\frac{0.6}{0.339/\sqrt{5}}=3.96.

Denne er t-fordelt med 4 frihetsgrader under H0.

Dette må sammenlignes med en passende tabellverdi. På 5 % nivå er dette 2,132. Altså kan vi forkaste H0, og påstå at mosjon reduserer kolesterolnivået.

1. Her får vi en to-utvalgs t-test fordi vi har uavhengige data (det er forskjellige folk). La M og D være effekt av henholdsvis mosjon eller diet med hensyn på reduksjon i kolesterol. Antar at standardavviket (s)er det samme for begge behandlinger.

H0: M = D  mot H1: M ≠ D  .

Forkast H0 hvis |T| > t 0,025, 9 = 2,262, der T =  = 0.43452

**Eller ved hjelp av R-comander**

Two Sample t-test

t = 0.43452, df = 9**, p-value = 0.3371**

alternative hypothesis: true difference in means is greater than 0

sample estimates:

mean of x mean of y pooled std.dev.

0.6000000 0.5000000 0.3800585

**Oppgave 7**

1. Parvis, det er samme kvinne.  Antall Migreneanfall uten og med medisin for en kvinne er **avhengige** variable. (Kvinnen kan være sterkt eller svakere disponert for migrene). Det blir derfor direkte feil å analysere dette som to uavhengige utvalg.

La Xivære antall årlige anfall uten Nembuxil (kvinne nr. i)., og

Yi være antall anfall for samme kvinne med bruk av Nembuxil (kvinne nr. i).

Vi lar Di være Xi - Yi.

Modell: Di ~ N(\mu_D, \sigma_D), der alle Di–ene er uavhengige. i = 1, 2, . . . , 9.

Estimert effekt av Nembuxil: \hat{\mu}_D=\bar{D}=2.9.  (\hat{\sigma}_D = sD = 3,3)

**b)**   H0: \mu_D = 0 (det vil si ingen effekt) mot  H1: \mu_D > 0 (det vil si positiv effekt)

T=\frac{\bar{D}}{s_D/\sqrt{9}}=\frac{2.9}{3.3/3}=2.64

Må sjekkes mot t-tabell med 8 frihetsgrader som er 1,86 ved 5 % nivå.

Dermed kan vi forkaste H0 og påstå av Nembuxil har positiv effekt.

**Oppgave 8**

Dette må være parvis test siden vi har med tvillinger å gjøre. Ser også av data at aggressivitet er framtredende hos noen tvillinger, mens det hos andre er langt mindre problem.

La Xivære aggressivitet hos førstefødte innen tvillingpar i, og

Yi være aggressivitet hos sistefødte innen tvillingpar i,

Vi lar Di være Xi - Yi.

Modell: Di ~ N(\mu_D, \sigma_D), der alle Di–ene er uavhengige. i = 1, 2, . . . , 8.

Tvillingpar Differanser

1. 6
2. 2
3. 0
4. -1
5. 1
6. -4
7. 4
8. 4

 = 1,5

SD = 3,2

95 % KI er gitt ved (  ) = (-1,18; 4,18).

Siden intervallet dekker 0, kan vi ikke si at psykologens påstand er bevist.

# Oppgave 9

**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

**Hvis det er like mange i hver gruppe, så har vi: **

**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.**

*1* Er effekt av fiskeolje med hensyn kolestrolreduksjon (i betydning gjennomsnittlig kolesterolreduksjon i befolkningen dersom alle tok fiskeolje).

Tilsvarende for *2*, men da matolje

**Spredning i kolesterolreduksjon for alle (hele populasjon)som bruker ***samme*** oljetype

Hvilken konklusjon trekker du?

Er sikker (på signifikansnivå 0,5 %) på at fiskeolje er bedre enn matolje dersom en ønsker reduksjon i kolesterol.