**Stat100 vår 2018**

**Løsning til uke 12**

**Oppgave 1**

La n = 9, og  = 11.

P(> 11| = 10) = P() = P(Z > 1) = 0,16. Behold H0

**----------------------------------------------------------------------------------------------**

La n = 16, og  = 11.

P(> 11|= 10) = P() = P(Z > 1,33) = 0,092. Forkast H0 på 10 % nivå, men ikke på 5 % nivå.

**-----------------------------------------------------------------------------------------------**

La n = 9, og  = 12

P(> 12|= 10) = P() = P(Z > 2) = 0,023. Forkast H0 på 5 % nivå, men ikke på 1 % nivå.

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

La n = 16, og  = 12.

P(> 11|= 10) = P() = P(Z > 2,73) = 0,0038. Forkast H0 på 1 % nivå.

P-verdien avtar når avstanden mellom  og  blir større.

P-verdien avtar med økende antall observasjoner (n).

Et utvalgssnitt på 12 er mer verdifullt dersom n = 16 i forhold til om n = 9.

**Oppgave 2**

1. *H0: u = 16,5 H1: u > 16,5.*

5 % nivå: Forkast H0 hvis Z = > 1,64.

Siden Z = 0,75, kan vi ikke forkaste H0, og må si til produsenten: Mulig du har rett, men Du har ikke klart å bevise din påstand.

P-verdien finner du ved P(Z > 0,75) = 0,23. Altså er det 23 % sjanse for at gjennomsnittet av 9 prøver blir større eller lik 17 selv om gjennomsnittet av alle produserte fiskepuddinger er på 16,5.

**Oppgave 3**

Anta at eggvekt for egg fra ny leverandør er normalfordelt med ukjent forventning **** og kjent standardavvik 10 gram.

**A)** H0: Ny leverandør gir ikke tyngre egg.  = 360

H1: Ny leverandør gir tyngre egg.  360.

Utvalgsgjennomsnittet () blir 366.4. gram

P-verdi finner du ved: P( > 366,4) = P(Z > ) = P(Z > 2,86) = 0,002,

der Z er standard normalfordelt.

Vi kan dermed forkaste H0 på alle rimelige signifikansnivå og konkludere med H1, ny leverandør er bedre (i betydning selger tyngre egg).

**B)** La Y være antallet kartonger som veier mer enn 360 gram.

Y er da Bin(20, p) der p = P(en kartong veier mer enn 360 gram).

Dersom mer enn halvparten veier mer enn 360 gram vil ny leverandør gi tyngre egg.

H0: Ny leverandør gir ikke tyngre egg. p =0,5

H1: Ny leverandør gir tyngre egg. p  0,5.

Vi finner 14 kartonger på mer enn 360 gram. **=** 14/20 = 0,7

P-verdi finner du ved:

P(**>** 0,7|p = 0,5) = P(**)** = P(Z > 1,78) = 0,0375,

**C)** Betraktelig større p-verdi i B, fordi vi der bare ser om en kartong er tung eller ikke tung, mens i A registrerer vi også hvor tung kartongen er.

Altså tar vi vare på mer informasjon i data i testen i A enn testen i B, og er derfor er testen i A smartere. Men testen i A antar normalfordeling, hvis dette ikke er sant er vi ute på ville veier.

**Oppgave 4**

**a)** H0: Nytt for gir ikke mindre spekk: ≥ 12

H1: Nytt for gir mindre spekk: < 12.

**b)** Type 1 feil. Påstå ***feilaktig*** at nytt for gir mindre spekk, og dermed bedre gris.

Type 2 feil: Ikke akseptere at nytt for er bra selv om det er det.

**c)** Regel: Forkast H0 på signifikansnivå  hvis Z < - z, der Z =  og z er avhengig av signifikansnivå. Vi har altså  < - z, da må < 12  z

Vi får for 9 griser:

Signifikansnivå 0,1: z = 1,28 dermed < 12 – 1,28\*= 11,15

Signifikansnivå 0,05: z = 1,64 dermed < 12 – 1,64\*= 10,9

Signifikansnivå 0,01: z = 2,33 dermed < 12 – 2,33\*= 10,45

**d)** Vi får for 81 griser:

Signifikansnivå 0,1: z = 1,28 dermed < 12 – 1,28\*2/9= 11,72

Signifikansnivå 0,05: z = 1,64 dermed < 12 – 1,64\*2/9= 11,64

Signifikansnivå 0,01: z = 2,33 dermed < 12 – 2,33\*2/9 = 11,48

**e)** Dersom H0 er sann har vi:

P-verdi er P( <11) = P( < ) = P(Z < ) =P(Z < )

ved n = 9: P(Z < ) = P(Z < -1,5 ) = 0,0668

ved n = 81: P(Z < ) = P(Z < -4,5 ) = 0,0000…

**f)** P-verdi er sannsynligheten for å få et gjennomsnitt i utvalget på 11 millimeter eller mindre dersom forventet verdi (gjennomsnitt i populasjonen) er 12 millimeter (H0 er sann)

Ved 9 griser kan vi forkaste H0 på 10 %, men ikke på 5 % nivå (og selvsagt ikke på 1 % nivå).

Ved 81 griser kan vi forkaste H0 på alle nivå.

**Oppgave 5**

La X være antall barn som har diaré ***etter*** rensing, da er X ~ B(100, p)

Der p er andelen barn (i hele populasjonen) som lider av diaré, ***etter*** at brønnen er renset.

Vi skal teste

H0: p = 0,6 (ingen effekt av rensing)

H1: p < 0,6 (positiv effekt av rensing)

Forkastningsområdet for testen blir ved 5 % nivå.

Z=  > 1,64, noe betyr at:

P-verdien på testen finner du slik:

P(< 0,5|p = 0,6)

P() = P(Z< -2,41) = 0,02

Vi forkaster på signifikans nvå mindre enn 0,02.

På 5 % nivå: Vi er sikre på at rensing har positiv effekt.

På 1 % nivå: Vi kan ikke påvise at rensing har positiv effekt.

**Oppgave 6** Utskrift fra R-commander

a) mean sd n

**4.76** 0.438 10

Estimerer all gås til å ha snittvekt på 4,76 kg.

Standardavviket til dette estimatet er 0,158 kg. (0,5/kvadratrot10)

1. Utskrift fra R-commander

99 percent confidence interval:

**4.45 5.07**

1. 4,5 ligger inne i 99 % KI. Vi Kan ikke forkaste H0 på 1 % signifikansnivå.
2. Utskrift fra R-commander

z = 1.6444, p-value = 0.1001

1. Hva slags konfidensintervall vil ikke dekke 4,5?

89 %

Utskrift fra R-commander

89 percent confidence interval:

4.507303 5.012697

Bonden er usikker på om han vil slakte nå, så han utvider utvalget av gjess til totalt å bli 32. **Gjennomsnittsvekten i utvalget er nå 4.7**

1. 99 percent confidence interval:

4.47 4.93

1. Vil du råde bonden til å slakte? Ja

**Oppgave** 7. ***Hastighet på biler***

1. = 60,4 km/t
2. Standardfeil er  der s er det estimerte standardavviket.

s= 9,66, dermed blir standardfeilen til lik 4,32

1. H0:  = H1:  > 52

P( ≥ 61,8| = ) = P(Z ≥ ) = P(Z ≥ 4.49) = 0,000000…..

Dette er P-verdien. Vi forkaster H0 og tror på H1.Gjennomsnittsfarta har gått opp.

**Oppgave 8***Seleksjon mot miljøgifter* **(Du må bruke R-commander )**

1. Sett opp en modell for forsøket.

La Yi være antall egg lagt av flue nr, i. i = 1,2, . . . .25

Vi antar at Yi -ene er uavhengige og Yi ~ N(, 60.)

Foretrekker modellen på denne formen:

La Yi være antall egg lagt av flue nr, i. i = 1,2, . . . .25

Yi =  +ei, der ei ~ N( , 60.)

1. mean

**205.48** Standardavviket til gjennomsnittet er 60/5 = 12.

1. Hvordan vil du estimere effekten av seleksjon med hensyn på fruktbarhet? (Forventet endring)

250 – 205,48 = *44,52*. Dette er vårt estimat på effekten av seleksjon på fruktbarhet.

Altså: Vi antar (estimerer) at ved seleksjon vil fruktbarheten reduseres i gjennomsnitt med ca. 44 egg pr flue.

H0:  = 

H1:  < 250

Utskrift fra R-commander

z = -3.71, p-value = 0.0001036

alternative hypothesis: true mean is less than 250

1. Forklar hva p-verdien sier i dette tilfelle. Trekk konklusjon på basis av p-verdi

P-verdien sier hvor sannsynlig er det å observere et snitt av 25 observasjoner på 205,48 eller mindre dersom forventningen er på 250. Vi ser at dette er nærmest umulig, som en konsekvens av dette forkaster vi null hypotesen og tror på alternativet, slik av vi er sikre på at seleksjon har negativ effekt på fruktbarhet.

1. P-verdien vil gå opp ved Feil 1 fordi avstanden mellom forventning og det observerte blir mindre.

P-verdien vil gå ned ved Feil 2 fordi det er mindre variasjon i naturen.

Ved begge feil vil de trekke i hver sin retning, og vi lar datamaskinen hjelpe oss til å se om P-verdien går opp eller ned.

Feil 1:

*One Sample z-test*

*data: flue$antall.egg.selektert*

*z = -2.0433, p-value =* ***0.02051***

*alternative hypothesis: true mean is less than* ***230***

*Feil 2*

*One Sample z-test*

*data: fluer$antall.egg.selektert*

*z = -5.565, p-value = 0.00000001311*

*alternative hypothesis: true mean is less than* ***250***

*Både Feil 1 og Feil 2:*

data: fluer$antall.egg.selektert

z = -3.065, , p-value = 0.001088

alternative hypothesis: true mean is less than **230**