**Stat100 vår 2018**

**Oppgaver til uke 12**

**Oppgave 1**

Anta du har et tilfeldig utvalg av n observasjoner fra en normalfordelt populasjon med ukjent forventning , men med kjent standardavvik; ** = 3**.

Finn P-verdien for følgende tester, samt bestem deg om du vil forkaste H0 eller beholde den.

n = 9. Test H0: = 10 mot H1: > 10 når  = 11

n = 16. Test H0: = 10 mot H1: > 10 når  = 11

n = 9. Test H0: = 10 mot H1: > 10 når  = 12

n = 16. Test H0: = 10 mot H1: > 10 når  = 12

Forklar hvordan disse faktorene påvirker P-verdien:

* Avstanden mellom () og forventningsverdien () under H0.
* Antallet observasjoner (n).

**Oppgave 2**

Fra tidligere er det kjent at standardavviket () i proteinprosent i fiskepudding er 2 (prosent). Anta at du produserer et stort parti fiskepudding og gjerne vil finne gjenomsnittlig proteininnhold () i **all** fiskepudding du har produsert. Du tar ut 9 prøver som gir et gjennomsnitt på 17 % protein.

Produsenten påstår at hans fiskepudding i snitt garantert inneholder minst 16,5 % protein. Gir du han rett i dette: Svar ved hypotesetesting, der du både bruker signifikansnivå 5 % og der du finner P-verdien.

**Oppgave 3 (Løses gjerne ved R-commander: Statitics > Means > Z-test > One sample Z (summarizes data)**

En butikksjef vurderer å bytte egg-leverandør.

Den forrige han brukte, leverte 6-pakninger med gjennomsnittsvekt (forventning) 360 gram og standardavvik 10 gram. En annen leverandør påstår at han har tyngre egg som han tilbyr til samme pris. Butikksjefen vil ikke bytte med mindre han er rimelig sikker på at ny leverandør virkelig har tyngre egg. Han kjøper 20 kartonger av den nye leverandøren, veier disse og får følgende:

348 352 355 357 357 358 361 362 362 362

365 370 371 374 376 376 379 380 381 382

Anta vekt fra ny leverandør er normalfordelt med ukjent forventning, men samme standardavvik som tidligere.

1. Ta utgangspunkt i oppgitte vekter. Utfør en hypotesetest for å sjekke om han skal bytte leverandør. Finn P-verdien til testen. Konkluder!
2. I sted for å ta utgangspunkt i vektene, tar butikksjefen og teller opp hvor mange av kartongenesom er tyngre enn 360 gram. Basert på dette tallet utfør en hypotesetest for å sjekke om han skal bytte.
3. Hva er fordelen svakheten ved å velge en av de to testene som er gjort over?

Hvorfor tror du at strategi 1 gir lavere p-verdi?

**Oppgave 4**

Anta at du har konstruert et nytt fór til gris, der hensikten er å redusere spekktjukkelsen (fettlaget) på grisen. Det er kjent at med dagens diett er gjennomsnittlig (forventet) spekktjukkelse 12 mm med et standardavvik på 2 mm. Vi ønsker å bevise at ditt fór gir mindre spekk. Vi antar at standardavviket ikke endrer seg med det nye fóret.

**a)** Sett opp nullhypotese og alternativet.

**b)** Hva er type 1 feil og type 2 feil i denne sitasjonen?

**c)** Hva blir maksimal gjennomsnittlig spekktjukkelse du kan ha for å forkaste nullhypotesen ved henholdsvis 10 %, 5 % og 1 % signifikansnivå med 9 griser?

**d)** Hva blir maksimal gjennomsnittlig spekktjukkelse du kan ha for å forkaste nullhypotesen ved henholdsvis 10 %, 5 % og 1 % signifikansnivå med 81 griser?

**e)** Dersom du fikk et gjennomsnittlig spekktjukkelse på 11 mm, hva blir P-verdien ved henholdsvis 9 og 81 griser?

**f)** Forklar hva P-verdien du fant i **e)** betyr og hvordan du kan bruke denne til å trekke konklusjoner på signifikansnivå henholdsvis 0,1, 0,05 og 0,1.

**Oppgave 5** *Har det positiv effekt å rense drikkevann?*

Anta at det er kjent at 60 % av barn i et fattig område led av diaré. Det ble bestemt at alle drikkevannsbrønnene skulle renses. Vi skal teste om rensing har positiv effekt med hensyn på diareforekomsten.

Sett opp nullhypotese og alternativ.

Etter brønnrensing, plukker vi tilfeldig ut 100 barn fra landsbyen, og teller opp hvor mange som har diaré. Det var 50 stykker.

Regn ut testens P-verdi, og gi en forklaring på hva denne måler.

Siden rensing er svært dyrt vil i vi ikke påstå at det har positiv effekt med mindre vi er rimelig sikre på dette.

Hvilken konklusjon trekker du basert på P-verdi dersom du tester på 5 % eller på 1 %?

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

**Sammenhengen mellom testing og konfidensintervall er slik:**

Anta en tosidig test.

H0:  = 0H1:  ≠ 0

**Regel:**

Forkast H0 på signifikansnivå  dersom et (1 - )100% konfidensintervall ikke dekker 0.

Denne regel er helt generell og kan brukes for test av enhver parameter.

**Oppgave 6 Bruk gjerneR-commander: Statistics > Means > z-test > one sample z-test**

Kjøttkvaliteten på gås er best når gåsa veier mellom 4 og 5 kilo. *En oppdretter vil derfor slakte gjessene sine dersom han er sikker på at gjennomsnittsvekta ligger rundt 4.5 kilo.* Han tar et tilfeldig utvalg på 10 gjess og veier disse, noe som ga:

Vekt i kilo (X) : 4.7, 4.2, 4.9, 5.1, 4.0, 5.5, 4.9, 4.5, 4.8, 5.0

Tilleggsopplysninger: Vi antar at slaktevekter er normalfordelt og med kjent standardavvik = 0,5.

1. Estimer gjennomsnittsvekta for all gås denne oppdretteren har og gi standardavviket til denne estimatoren.
2. Lag et 99 % konfidensintervall for forventet gåsevekt.
3. Bruk intervallet til å teste den **tosidige** testen der H0:  = 4,5.

Hva blir nivået på testen?

Kan du forkaste H0?

1. Finn testens p-verdi.
2. Hva slags konfidensintervall vil ikke dekke 4,5? Hint bruk p-verdien på den tosidige testen.

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. Hva blir et 99% konfidensintervall for  i dette tilfellet?
2. Vil du råde bonden til å slakte?

**Oppgave** 7. ***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 og et standardavvik på 10km/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

1. Hva er et forventningsrett estimat på gjennomsnittlig (forventet) fart etter fartshevningen?
2. Hva er standardfeilen til dette estimatet (anta at standardavviket kan ha endret seg med fartshevingen)?
3. Du får i oppgave å teste om hevingen fra 50 til 60 km/t har ført til en forventet økning i fart på strekningen og synes 5 biler er for lite i utvalget. Du måler derfor farten til 16 nye biler og finner at gjennomsnittsfarten i det nye utvalget er 61.8 (Dermed har du 21 observasjoner). Anta at standardavviket fortsatt er 10 km/t (= 10).

Sett opp hypotesene, finn P-verdien og konkluder.

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

Bananfluer ble selektert for motstand ovenfor miljøgifter brukt i landbruket. For disse talte man opp antall egg lagt av hunnfluer i løpet av 14 dager tidlig i deres livssyklus. Vi ønsker å undersøker effekten av denne seleksjonen med hensyn på fruktbarhet. Vi antar at seleksjon ikke påvirker variasjoner i fruktbarhet. Vi vet på forhand at forventet antall egg for ikke-selekterte er 250 med et standardavvik  = 60.

Data (bananfluer) ligger på canvas under Modul Data.

Last ned fila Banafluer (R.datafil) til egen maskin.

For å lese data inn i R-commander: DATA > LOAD DATA.

1. Sett opp en modell for forsøket.
2. Estimer forventet antall egg og gi standardavviket til estimatet.
3. Hvordan vil du estimere effekten av seleksjon med hensyn på fruktbarhet? (Forventet endring)
4. Test om det er en sikker (signifikant) negativ effekt på fruktbarhet ved å selektere på giftmotstand. Sett opp både nullhypotesen og den alternative hypotesen.
5. Forklar hva p-verdien sier i dette tilfelle. Trekk konklusjon på basis av p-verdien.
6. Anta at det var oppgitt 2 feil i teksten over

1: Forventet antall egg for ikke-selekterte var 230.

2: Standardavviket er 40.

Gjenta hypotesetestingen når du retter opp:

Bare Feil 1

Bare Feil 2

Både Feil 1 og Feil 2.

Kommenter hvordan det går med p-verdiene i de tre tilfellene sammenlignet med den du fant i E).