I oppgaver der det eksplisitt står at en skal vise mellomregninger, vil korrekt svar ikke bli godkjent dersom mellomregninger mangler.

Det brukes desimalpunktum i dette oppgavesettet.

Oppgavesettet består av 10 delpunkter som alle teller likt ved bedømming.

Oppgave 1

Innen luftfart er det satt krav til at en bestemt type bolt skal ha en strekkstyrke (X) på minst 18 kN.

a) Anta $X \sim N(19.0,\ 0.4)$, dvs. $\mu = 19.0,\ \sigma = 0.4$.

- Hva er sannsynligheten for at en tilfeldig bolt ikke oppnår kravet til strekkstyrke? Vis mellomregninger.

  PS. Mangler du svar her og har bruk for det senere i oppgavesettet, da kan du benytte 0.01 (ikke det korrekte svaret på denne oppgaven).

  Løsningsforslag: $P(X < 18.0) = G(\frac{18-19}{0.4}) = G(-2.5) = 0.0062$

- En komponent er festet med to bolter.
  Hva er sannsynligheten for at begge bolter er svakere enn kravet til strekkstyrke? Du kan anta boltene er uavhengige mht. strekkstyrke.

  Løsningsforslag:
  $P(X_1 < 18.0) \cdot P(X_2 < 18.0) = (0.0062)^2 = 0.00003844$

  Eventuelt dersom en ikke har funnet svaret i forrige punkt og må bruke den oppgitte verdien $p = 0.01$, blir svaret her $0.01^2 = 0.0001$,

Selskapet som mottar boltene bestemmer seg for å undersøke strekkstyrken, én bolt om gangen. La Y være antall forsøk de må gjøre før de finner den første bolten som ikke tåler strekkstyrke 18 kN.
Antagelsen $X \sim N(19.0, 0.4)$ gjelder forsatt.

b)
- Hvilken fordeling har Y.
  Løsningsforslag: Geometrisk fordeling med $p = 0.0062$ fra 1a).
- Hvor mange bolter forventes det å måtte undersøke før en finner den første som ikke tåler 18 kN?
  Løsningsforslag: $E(Y) = \frac{1}{p} = \frac{1}{0.0062} = 161.3$
  Med "nødløsningen" $p = 0.01$ fra a) er svaret $\frac{1}{0.01} = 100$

Selskapet som mottar boltene er ikke sikker på at den oppgitte forventede strekkstyrken $E(X) = \mu = 19.0$ stemmer. De velger ut 100 tilfeldige bolter og registrer strekkstyrke (X) til hver av disse.

c) Resultatene av de fem første boltene er: 19.1 19.1 19.5 18.6 18.2 kN

- Finn gjennomsnittet og standardavviket av disse fem målingene.
  Løsningsforslag: $\bar{X} = \frac{19.1+19.1+19.5+18.6+18.2}{5} = \frac{94.5}{5} = 18.9$ og
  $s = \sqrt{\frac{(19.1-18.9)^2+(19.1-18.9)^2+(19.5-18.9)^2+(18.6-18.9)^2+(18.2-18.9)^2}{5-1}} = \sqrt{\frac{1.02}{4}} = 0.50$

Resultatet fra alle 100 målinger er $\bar{X} = 18.85, \quad s = 0.42$.

d)
- Finn estimater for μ og σ basert på alle 100 målinger.
  Løsningsforslag: Estimat for μ er $\hat{\mu} = \bar{X} = 18.85$ og estimat for σ er $s = 0.42$
- Undersøk med hypotesetest om forventet strekkstyrke er mindre enn 19 kN. Bruk 5% signifikansnivå.
  -Formulér hypoteser.
  -Gjennomfør testen og konkludér.
  Du kan bruke følgende resultat fra R-funksjonen t.test():
  $T = -3.55$ og p-verdi$= 0.0003$ for ensidig alternativ hypotese.
  Løsningsforslag:
  $H_0 : \mu \geq 19$ versus $H_1 : \mu < 19$
  Den korrekte testen her er en T-test med ensidig alternativ hypotese. T-test fordi σ er ukjent og estimeres med s.
  Dersom vi bruker de gitte opplysningene, kan en gå direkte til konklusjonen og forkastes H_0 på 5% nivå fordi p-verdien (gitt lik 0.0003) er mindre enn 0.05.
  Uten bruk av de gitte opplysningene må vi først finne verdien av testobservatoren: $T = \frac{18.85-19.0}{\frac{0.42}{\sqrt{100}}} = -3.57$

som har $n-1=100-1=99$ frihetsgrader.
Kritisk verdi for denne ensidige t-testen med 99 frihetsgrader (leser av for 100 frihetsgrader i tabell) er $-t_{0.05}=-1.660$.

Konklusjon: Nullhypotesen forkastes fordi $T=-3.57<-1.660$.

En kan alternativt finne p-verdien til resultatet $T=-3.57$.
Det nærmeste en kommer med tabellen i læreboka er
$P(T<-2.626)=0.005$.
Da vet vi at p-verdien $=P(t<-3.57)<0.005$.

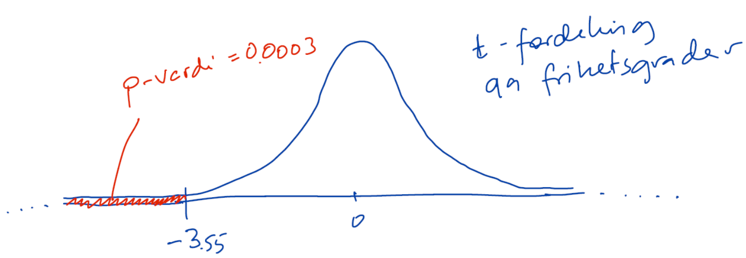
En fjerde variant på hypotesetesten er å finne forkastingsområdet for $\bar{X}$: Forkast H_0 dersom $\bar{X}\le 19.0-1.660\cdot\frac{0.42}{10}=18.93$,
Den observerte verdien $\bar{X}=18.85$ er mindre enn 18.93 og nullhypotesen forkastes.

Konklusjon: p-verdi er mindre enn signifikansnivået og nullhypotesen forkastes. Forventet strekkstyrke er mindre enn 19 kN.

Alle varianter her av hypotesetesten er identiske og likeverdige.

e) Lag en passende skisse og markér p-verdien i skissen.

Løsningsforslag:



f) Lag et 95% KI for σ og bruk intervallet til å si noe om modellantagelsen $\sigma=0.4$.

Løsningsforslag:
Nedre grense i KI for σ^2: $\frac{(n-1)\cdot s^2}{\chi_{0.025}}=\frac{99\cdot 0.42^2}{129.56}=\frac{17.4636}{129.56}=0.1348$

Øvre grense i KI for σ^2: $\frac{(n-1)\cdot s^2}{\chi_{0.975}}=\frac{99\cdot 0.42^2}{74.22}=0.2352$

Et 95% konfidensintervall for σ er da lik:
$\left(\sqrt{0.1348},\quad \sqrt{0.2352}\right)$
Dvs. intervallet $(0.367,\quad 0.485)$

PS. Skulle egentlig hatt tabellverdier for 99 frihetsgrader. Her er nærmeste verdi brukt, som er for 100 frihetsgrader.

I tillegg til gjennomsnittet $\bar{X} = 18.85$ og standardavviket $s = 0.42$ fra alle $n = 100$ målinger, oppsummerer vi enkeltmålingene i fem intervaller:

Intervaller	"< 18.49"	"18.49 - 18.74"	"18.74 - 18.96"	"18.96 - 19.21"	"$19.21 >$ "
Antall bolter	17	26	20	22	15

Her er intervallgrensene laget slik at en normalfordeling $N(18.85,\ 0.42)$ har omtrent like stort areal i hvert av de fem intervallene.

g) Undersøk med en modelltest om dataene er observasjoner fra normalfordeling med forventningsverdi 18.85 og standardavvik 0.42

- Vis at forventet antall i hvert av de fem intervallene i frekvenstabellen over er ca 20, dvs. det holder her å vise at dette er tilfelle for det første intervallet "< 18.49".

  Løsningsforslag:
  $E_1 = P(X \leq 18.49) = G(\frac{18.49-18.85}{0.42}) = G(\frac{-0.36}{0.42}) = G(-0.86) = 0.1949 \approx 0.2$

- Regn ut testobservator. Her kan du bruke at alle $E_i = 20$.

- Angi antall frihetsgrader og konkludér i hypotesetesten.

Løsningsforslag:

Intervaller	< 18.49	18.49 - 18.74	18.74 - 18.96	18.96 - 19.21	$19.21 >$
Observert antall (O)	17	26	20	22	15
Forventet antall (E)	20	20	20	20	20
Avvikssum $\frac{(O-E)^2}{E}$	$\frac{9}{20}$	$\frac{36}{20}$	0	$\frac{4}{20}$	$\frac{25}{20}$

$Q = \frac{74}{20} = 3.7$

Med $k = 5$ kategorier har vi $k - 1 = 4$ frihetsgrader.
Testobservator Q er tilnærmet kji-kvadratfordelt med 4 frihetsgrader. Kritisk verdi for 5% signifikansnivå er 9.49. Den observerte Q-verdien er ikke stor nok til å forkaste nullhypotesen om normalfordeling. Det er ikke noe i datamaterialet som tyder på at målingene ikke kommer fra en normalfordelt populasjon.

Oppgave 2

Bolter produsert til et bestemt formål må være av en bestemt lengde. Bolter som er for lang kan kuttes til riktig lengde, men bolter som er for korte må kasseres. Produsenten opplyser at 89% av boltene har riktig lengde, 6% er for lange og 5% for korte. Dvs. at 95% av boltene kan benyttest.
Boltenes lengder er uavhengige.

a)
- Finn sannsynligheten for at minst 9 i et utvalg av 10 bolter kan brukes, enten direkte eller etter kutting.

  Løsningsforslag: $Y \sim Bino(10, 0.95)$ der $p = 0.89 + 0.06 = 0.95$
  Bruker tabell E.1 i læreboka og finner
  $P(Y \geq 9) = 1 - P(Y \leq 8) = 1 - 0.086 = 0.914$

- La X være antall bolter i et utvalg av 20, som ikke kan benyttes fordi de er for korte.

  Bruk en passende tilnærmingsfordeling og finn sannsynlighetene:
  $P(X = 0)$
  $P(X \leq 3)$

  Løsningsforslag:
  Har at $X \sim Bino(20, 0.05)$
  Benytter tilnærming til Poissonfordeling (Regel 5.9) fordi $n = 20 > 10$ og $p = 0.05 < 0.1$. Fra tabell E.2 finner vi med $\lambda = 20 \cdot 0.05 = 1.0$
  $P(X = 0) = 0.368$
  $P(X \leq 3) = 0.987$
  OBS, må løses med Poissonfordelingen fordi oppgaven går ut på å nettopp finne en passende tilnærmingsfordeling.

Produsenten er fornøyd med at kun 5 % av boltene er for korte og må kasseres. Produksjonsutstyret er gammelt og en mistenker at denne andelen har økt. Produsenten har i et utvalg av $n = 200$ bolter funnet at 15 bolter var for korte og ikke kunne brukes.

b) Undersøk med hypotesetest om andel 'for korte' bolter har økt.

- Formuler hypoteser.

  Løsningsforslag: $H_0 : p = 0.05 = p_0$ mot $H_1 : p > 0.05$
  eller ekvivalent $H_0 : p \leq 0.05$ mot $H_1 : p > 0.05$

- Regn ut testobservatoren.

Løsningsforslag: Bruker tilnærming til normalfordeling (Regel 5.20) fordi $np_0(1-p_0) = 200 \cdot 0.09 \cdot 0.95 = 9.5 > 5$
Testobservator $Z = \frac{X-200\cdot 0.05}{\sqrt{200\cdot 0.05\cdot 0.95}} = \frac{5}{\sqrt{9.5}} = 1.62$

- Skriv en konklusjon av testen. Bruk 5% signifikansnivå.

  Løsningsforslag:

  Kritisk verdi for forkasting av nullhypotesen er $z_{0.05} = 1.645$

  Testobservatoren er ikke større enn 1.645 og kan dermed ikke forkaste nullhypotesen på 5% nivå.

c)
- Hvor stor er sannsynligheten for å gjøre en type I feil i denne testen?
- Hva kalles denne sannsyligheten?
- Finn p-verdien til datamaterialet for denne hypotesetesten.

Løsningsforslag:

Type I feil, er det å forkaste en sann nullhypotese. Her er denne sannsynligheten bestemt til å være 5% ($\alpha = 0.05$)

Sannsynligheten for å gjøre en type I feil er dermed 0.05

Sannsynligheten kalles signifikansnivået til testen.

p-verdien= $P(Z \geq 1.62) = 1 - G(1.62) = 1 - 0.9474 = 0.0526$