Høgskolen i Østfold

EKSAMEN

Emnekode:	Emnenavn:
IRB22515, IRBIO22013, IRE22518, IRM23116	**Deleksamen i Statistikk**
Dato: **03.01.19** **Sensurfrist:** **24.01.19**	**Eksamenstid:** **09.00 – 12.00**
Antall oppgavesider: 6 **Antall vedleggsider:** 9	**Faglærer:** Tore August Kro **Oppgaven er kontrollert:** Ja

Hjelpemidler:

Lærebok, to interne notater (begge Elise Øby 2015),

kalkulator av enhver type, godkjente formelsamlinger.

Om eksamensoppgaven:

Gjør alle oppgavene. Alle deloppgaver teller likt.

Vis alle utregninger.

Besvarelsen vurderes ut fra kvaliteten på begrunnelsene.

Kandidaten må selv kontrollere at oppgavesettet er fullstendig

Høgskolen i Østfold
Avdeling for ingeniørfag

EKSAMEN
STATISTIKK **Lærer/telefon: Tore A. Kro, 900 22 321**

Deleksamen i Statistikk IRB22515, IRBIO22013 IRE22518, IRM23116	**Dato: 03.01.2019**	**Tid: 0900–1200**
Antall oppgavesider: 6	**Vedlegg: Ett internt notat, 9 sider, (Elise Øby, 2015)**	
Sensurfrist: 24.01.2019		
Hjelpemidler: Lærebok, to interne notater (begge Elise Øby 2015), kalkulator av enhver type, godkjente formelsamlinger.		
KANDIDATEN MÅ SELV KONTROLLERE AT OPPGAVESETTET ER FULLSTENDIG		

Vis alle utregninger. Besvarelsen vurderes ut fra kvaliteten på begrunnelsene.

Oppgave 1. Sannsynligheten for at en tannlege som jobber i Østfold, ofte spiser svensk godteri er $p = 0{,}3$. Hos skoletannlegen i Fredrikstad jobber det 10 tannleger. La X være antall av disse som har denne uvanen.

a) Hvilken fordeling har X? Finn $E(X)$ og $\mathrm{Var}(X)$. Hva er sannsynligheten $P(X = 3)$?

Totalt jobber det 253 tannleger i Østfold. La Y være antall av disse som ofte spiser svensk godteri.

b) Er Y tilnærmet normalfordelt? Finn et tosidig 95% spredningsintervall for antall tannleger i Østfold som ofte spiser svensk smågodt.

Det er mange fra Østfold som shopper i Strömstad. Blant annet er svensk godteri billig. Dette påvirker tannhelsen. Dersom man unngår å spise svensk godteri er sannsynligheten for å oppdage hull i tennene på neste årskontroll hos tannlegen 0,05. Derimot er sannsynligheten for hull lik 0,35 blant personer som ofte spiser svensk godteri.

c) Den 3.januar går en tilfeldig tannlege fra Østfold til årlig kontroll. Vedkommende tannlege har hull i sine egne tenner. Hva er sannsynligheten for at denne tannlegen ofte spiser svensk godteri?

Oppgave 2. En prosess er poissonfordelt med rate $\lambda = 2$ hendelser pr minutt. La X være antall forekomster i løpet av $t = 3$ minutter.

a) Finn $E(X)$, $\mathrm{Var}(X)$ og $P(X \geq 2)$.

I neste deloppgave observerer vi en annen poissonprosess med en ukjent rate λ. Dersom man observerer prosessen en stund kan man estimere raten ved

$$\hat{\lambda} = \frac{\text{antall hendelser}}{\text{observasjonstid}}.$$

Anta at vi har observert prosessen så lenge at vi med rimelig sikkerhet har $\hat{\lambda} \approx 2$ hendelser pr minutt.

b) Dersom vi observerer prosessen over et enda lengre tidsrom t kan vi få et enda bedre estimat for raten. Anslå hvor stor t minst må være for å få et 95% konfidensintervall for raten λ som har bredde mindre enn 0,2.

En skoleklasse får et gruppeprosjekt i statistikk. Hver gruppe skal foreslå en praktisk situasjon som best mulig lager poissonfordelte hendelser med raten $\lambda = 2$ hendelser pr minutt. Her er forslagene til noen av gruppene:

Gruppe A: Elevene i denne gruppen mener at de som en samlet gruppe kan småjogge med en jevn fart på 6 km/t. De lager derfor en 50 m lang rundløype på en grusbane. Hendelsen de skal telle er antall passeringer av startstreken.

Gruppe B: Elevene i denne gruppen ønsker å kaste en vanlig sekssidet terning og telle antall ganger de slår en sekser. De planlegger å gjøre 12 kast pr minutt.

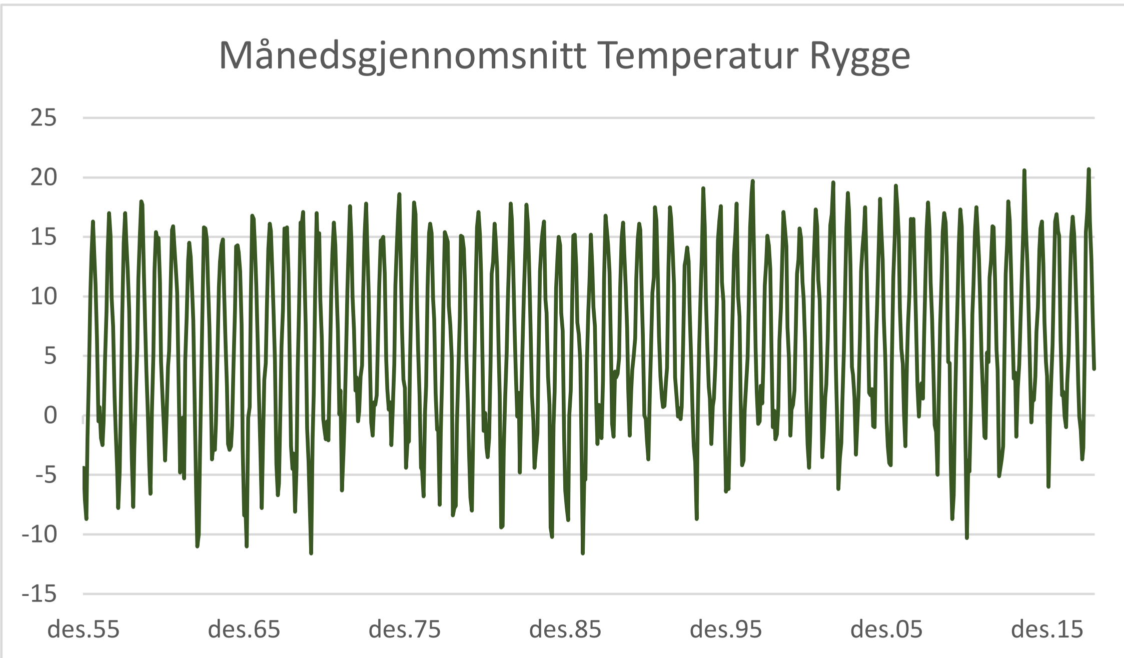
Gruppe C: Denne gruppen ønsker å fylle skolens svømmebasseng med levende laks. De seks elevene i gruppen skal deretter fiske med stang. Hendelsen, som skal telles, er antall ganger de får fisk på kroken. Laksen skal ikke slippes ut igjen etter den er tatt. De har funnet ut at med en laks i bassenget så er sannsynligheten 0,02 for at en elev fisker denne i løpet av en time. De starter med 1000 laks i bassenget.

Gruppe D: Elevene i denne gruppen vil benytte kortstokk. Hver elev skal ha en kortstokk. Etter tur skal en og en elev være i aksjon. Den aktive eleven skal rolig bla opp ett og ett kort fra en kortstokk og stoppe når spar ess kommer. Deretter skal neste elev i aksjon og denne eleven skal starte umiddelbart. Den neste eleven skal også bla opp kort på samme vis. Hendelsen som telles er antall spar ess. Elevene stokker kortene etter de har vært i aksjon, og hver elev kan være i aksjon mange ganger. Tempoet de blar opp kort med er viktig. Det skal gjøres så raskt at elevene ville brukt nøyaktig ett minutt på hele kortstokken dersom de ikke hadde stoppet på spar ess.

c) Ingen av gruppene har en eksakt poissonprosess. Gi en kort kommentar til hver gruppe på deres mest vesentlige mangel sammenlignet med en poissonprosess med rate $\lambda = 2$ hendelser på minutt.

Oppgave 3. I denne oppgaven skal vi bruke lineær regresjon for å undersøke om det finnes en trend i månedlige middeltemperaturer hentet fra Rygge målestasjon i tidsrommet fra desember 1955 til november 2018. Disse dataene kan man laste ned fra nettsiden `eKlima.met.no`.

Vi plotter månedlige middeltemperaturer fra Rygge målestasjon og får:



I datasettet er y-verdiene månedlig middeltemperatur målt i °C. Måleenheten for x-verdiene er år. Måneden januar 1955 lagres som $x = 1955$, mens november 2018 får verdien $x = 2018 + \frac{11-1}{12} \approx 2018{,}83$. På disse dataene utfører vi en lineær regresjon på signifikansnivå $\alpha = 0{,}05$. Dette er utskriften fra Excel:

SAMMENDRAG (UTDATA)

Regresjonsstatistikk	
Multippel R	0,09889245
R-kvadrat	0,00977972
Justert R-kvadrat	0,00846643
Standardfeil	7,45035051
Observasjoner	756

Variansanalyse

	fg	*SK*	*GK*	*F*	*Signifkans-F*
Regresjon	1	413,351191	413,351191	7,44673302	0,006503229
Residualer	754	41852,8229	55,5077227		
Totalt	755	42266,1741			

	Koeffisienter	*Standardfeil*	*t-Stat*	*P-verdi*	*Nederste 95%*	*Øverste 95%*
Skjæringspunkt	-74,588462	29,6117631	-2,5188795	0,01197868	-132,7197647	-16,45715999
År	0,04065829	0,01489931	2,72887028	0,00650323	0,011409225	0,069907363

a) Angi formelen for regresjonslinja. Bruk regresjonslinja til å anslå hvor mye temperaturen har endret seg fra 1957 til 2017.

b) Hva betyr det når et resultat er "signifikant"? I regresjonsanalysen over er det utført en hypotesetest. Formuler de aktuelle hypotesene i dette tilfellet. Hva blir konklusjonen?

Oppgave 4. Slankekongen Kong Slankeland produserer slankepulveret Lurium. For å dokumentere den fantastiske virkningen av dette slankepulveret bestiller Kong Slankeland fire uavhengige undersøkelser fra velvillige oppdragsforskere.

I hver undersøkelse skal Lurium gis til et representativt utvalg på 50 personer. Resultatet, som skal rapporteres videre, er den gjennomsnittlige vektreduksjonen blant deltagerne. La den stokastiske variabelen X være resultatet fra en slik oppdragsforskningsrapport.

Resultatene fra de fire undersøkelsene var:

5,4	5,5	5,6	6,7

a) Finn gjennomsnittet og standardavvik for dette datasettet.

Dersom alle oppdragsforskerne har utført bestillingsforskningen korrekt og samvittighetsfullt, vil det følge av sentralgrenseteoremet at X er normalfordelt.

b) Kong Slankeland ønsker å bevise at Lurium er et fantastisk slankepulver. Utfør derfor en hypotesetest på signifikansnivå $\alpha = 0{,}05$ der en av hypotesene er:

Forventet vektreduksjon ved bruk av Lurium er mer enn 5,0 kg.

Slankekongen mistenker at noen av oppdragsforskerne jukser. Det kan være at det pyntes på resultatene, eller at det er fabrikkert data. En måte å undersøke denne mistanken, er å sammenligne den empiriske fordelingsfunksjonen til X med normalfordelingens fordelingsfunksjon.

Den empiriske fordelingsfunksjonen er gitt ved

$$\hat{F}(x) = \frac{\text{antall observasjoner mindre eller lik } x}{n}.$$

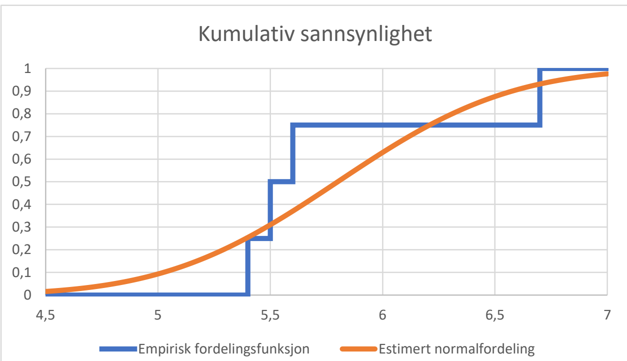
Normalfordelingens fordelingsfunksjon er gitt ved

$$F(x) = P(X \leq x) = G\left(\frac{x-\mu}{\sigma}\right),$$

hvor G er Gaussfunksjonen for kumulativ standardnormalfordeling.

c) Bruk de estimerte verdiene for forventningsverdi og standardavvik fra deloppgave a), og regn ut verdien $F(5{,}5)$ til normalfordelingens fordelingsfunksjon.
Forklar også hvorfor empirisk fordelingsfunksjon har $\hat{F}(5{,}5) = \frac{2}{4}$.
Finn differansen $F(5{,}5) - \hat{F}(5{,}5)$.

Følgende plott viser empirisk fordelingsfunksjon og normalfordelingens fordelingsfunksjon i samme koordinatsystem.



I deloppgave d) under skal du utføre Lilliefors-testen for normalfordeling. Testobservatoren i denne testen er tallverdien av maksimal differanse mellom disse funksjonene. Først må man sortere verdiene i datasettet i stigende rekkefølge, altså $x_1 < x_2 < \cdots < x_n$. Når man skal finne denne testobservatoren, regner man ut for hver observasjon x_i disse kritiske verdiene:

$$\left| G\left(\frac{x_i - \hat{\mu}}{S}\right) - \frac{i}{n} \right| \qquad \text{og} \qquad \left| G\left(\frac{x_i - \hat{\mu}}{S}\right) - \frac{i-1}{n} \right|.$$

Her er G Gaussfunksjonen, $\hat{\mu}$ er gjennomsnittet og S er utvalgets standardavvik. Testobservatoren D for Lilliefors-testen er den største av disse kritiske verdiene.

Du skal slippe å gjøre denne utregningen selv. I datasettet fra Kong Slankelands bestillingsforskning er det $n = 4$ observasjoner. Hver observasjon gir to kritiske verdier. Dette gir disse 8 mulighetene for maksimalverdien av avviket mellom empirisk fordelingsfunksjon og normalfordeling:

0,2544	0,0601	0,1294	0,1814
0,0044	0,1899	0,3794	0,0686

d) Under følger en beskrivelse av *Lilliefors-testen for normalfordeling.* Les beskrivelsen av denne hypotesetesten. Skriv setninger som forklarer hypotesene. Gjennomfør hypotesetesten på signifikansnivå $\alpha = 0{,}05$, og trekk en konklusjon.

Lilliefors-testen for normalfordeling.
La X være en stokastisk variabel. Gjør n målinger og sorter dem i stigende rekkefølge

$$x_1 < x_2 < \cdots < x_n.$$

Sett opp disse hypotesene:

H_0: X er normalfordelt $\qquad$ H_1: X er ikke normalfordelt

Testobservatoren D er maksimalt avvik mellom den empiriske fordelingsfunksjonen $\hat{F}$ og normalfordelingens fordelingsfunksjon F tilsvarende forventningsverdi lik gjennomsnittet $\bar{X}$ og standardavvik lik utvalgsstandardavviket S.

$$D = \max \left| \hat{F}(x) - F(x) \right|$$

Nullhypotesen skal forkastes dersom D er større enn den kritiske verdien $D_{n,\alpha}$. Disse kritiske verdiene er gitt i følgende utdrag av kvantiltabellen for Lilliefors-fordelingen.

Antall målinger		Areal	α	
n	0,01	0,05	0,10	0,20
4	0,4129	0,3754	0,3456	0,3027
5	0,3959	0,3427	0,3188	0,2893
6	0,3728	0,3245	0,2982	0,2694
7	0,3504	0,3041	0,2802	0,2521
8	0,3331	0,2875	0,2649	0,2387
9	0,3162	0,2744	0,2522	0,2273
10	0,3037	0,2616	0,2410	0,2171
11	0,2905	0,2506	0,2306	0,2080
12	0,2812	0,2426	0,2228	0,2004

1 Fordelinger og tilnærminger

1.1 Binomisk fordeling

En forsøksrekke består av n forsøk. Hvert forsøk har to mulige utfall: Suksess eller ikke suksess. Sannsynligheten for suksess er p i hvert forsøk. Variabelen X=*antall suksesser i løpet av n forsøk* er da binomisk fordelt og

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Forventning: $\mu = E(X) = np$. Varians: $\sigma^2 = \text{Var}(X) = np(1-p)$.

Tilnærming til normalfordelingen: For $\sigma^2 \geq 5$ er X tilnærmet normalfordelt: $X \simeq N(np, \sqrt{np(1-p)})$.

1.2 Hypergeometrisk fordeling

I en populasjon på N elementer har M elementer en spesiell egenskap. Det gjøres et utvalg på n elementer fra populasjonen. Variabelen X=*antall elementer med spesiell egenskap blant de n utvalgte elementene* er hypergeometrisk fordelt og

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Forventning: $\mu = E(X) = np$. Varians: $\sigma^2 = \text{Var}(X) = np(1-p)\frac{N-n}{N-1}$ der $p = \frac{M}{N}$.

Tilnærming til binomisk fordeling: Når $N >> n$ (hovedregel $N > 10n$) er X tilnærmet binomisk fordelt med suksessannsynlighet $p = \frac{M}{N}$.

Tilnærming til normalfordelingen: Når $\sigma^2 \geq 5$ er X tilnærmet normalfordelt: $X \simeq N\left(np, \sqrt{np(1-p)\frac{N-n}{N-1}}\right)$.

1.3 Poissonfordelingen

Antall forekomster av hendelsen A er Poissonfordelt hvis

(1) Antall forekomster av A i disjunkte tidsintervaller er uavhengige av hverandre

(2) Forventet antall forekomster av A er konstant lik λ per tidsenhet

(3) To forekomster av A kan ikke være fullstendig sammenfallende på tidsaksen

I løpet av de neste t tidsenhetene vil vi observere X forekomster av hendelsen A. Hvis Poissonforutsetningene er oppfylt, er X Poissonfordelt og

$$P(X = x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t}$$

Forventning: $\mu = E(X) = \lambda t$. Varians: $\sigma^2 = \text{Var}(X) = \lambda t$.

Tilnærming til normalfordelingen: Når $\sigma^2 = \lambda t \geq 10$ er X tilnærmet normalfordelt: $X \simeq N(\lambda t, \sqrt{\lambda t})$.

2 Sentralgrenseteoremet

2.1 Gjennomsnitt av variabler fra samme sannsynlighetsfordeling

La $X_1, \ldots, X_n$ $(n \geq 20)$ være uavhengige variabler **fra samme sannsynlighetsfordeling** med forventning μ og standardavvik σ. Da er

$$\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n) \simeq N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

2.2 Sum av variabler fra samme sannsynlighetsfordeling

La $X_1, \ldots, X_n$ $(n \geq 20)$ være uavhengige variabler **fra samme sannsynlighetsfordeling** med forventning μ og standardavvik σ. Da er

$$X_1 + \cdots + X_n \simeq N\left(n\mu, \sqrt{n}\sigma\right)$$

2.3 Sum av normalfordelte variabler

Hvis $X_1, X_2, \ldots, X_n$ er uavhengige og **normalfordelte** variabler med forventninger μ_i og varianser σ_i^2 der $i = 1, \ldots, n$, vil enhever sum av dem også være normalfordelt:

$$Y = a_1 X_1 + \cdots + a_n X_n$$

er normalfordelt med forventning

$$\mu_Y = a_1\mu_1 + \cdots + a_n\mu_n$$

og varians

$$\sigma_Y^2 = a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2$$

3 Estimering

3.1 Estimering av forventningsverdien μ når σ er kjent

La $X_1, X_2, \ldots, X_n$ være variable fra samme fordeling. Alle har forventningsverdi μ (som er den som er ukjent, og som vi skal estimere en verdi for) og kjent standardavvik σ. Vår beste gjetning for forventningsverdien er gjennomsnittet

$$\hat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Ved å gjøre et nytt utvalg av n variable fra denne fordelingen, vil vi få et nytt gjennomsnitt. Dermed kan vi se på $\overline{X}$ som en variabel i seg selv. Sentralgrenseteoremet gir at $\overline{X}$ er tilnærmet normalfordelt

$$\overline{X} \simeq N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Da er det f.eks 95% sikkert at en verdi $\overline{X}$ ligger i intervallet $\mu \pm 1,96 \cdot \frac{\sigma}{\sqrt{n}}$, som gir (ved å stokke litt om på ulikheter) at det er 95% sikkert at μ ligger i intervallet $\overline{X} \pm 1,96 \cdot \frac{\sigma}{\sqrt{n}}$. Dermed kan vi lage konfidensintervaller for den ukjente μ basert på en gjennomsnittsverdi:

$$\overline{X} \pm (z\text{-verdi som er bestemt av konfidensnivået}) \cdot \frac{\sigma}{\sqrt{n}}$$

3.2 Estimering av forventningsverdien μ når σ er ukjent

La $X_1, X_2, \ldots, X_n$ være variable fra samme fordeling. Alle har forventningsverdi μ (som er den som er ukjent, og som vi skal estimere en verdi for) og ukjent standardavvik σ. Vår beste gjetning for forventningsverdien er gjennomsnittet

$$\hat{\mu} = \overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Når σ er ukjent, må vi estimere denne også. Vår beste gjetning til variansen i populasjonen, er variansen i utvalget:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

Estimatet for σ blir da

$$\hat{\sigma} = S$$

Konfidensintervaller for μ med estimert verdi for σ lager vi slik:

$$\overline{X} \pm (\text{kritisk verdi fra } t\text{-tabellen med } (n-1) \text{ frihetsgrader}) \cdot \frac{S}{\sqrt{n}}$$

3.3 Estimering av sannsynlighet/andel p

La X være binomisk fordelt med suksess-sannsynlighet p (ukjent) eller hypergeometrisk fordelt med andel elementer i populasjonen med bestemt egenskap lik $\frac{M}{N} = p$. Ved å gjøre et utvalg på n forsøk og undersøke antall suksesser i forsøksrekken, kan vi beregne en estimert verdi for suksesssannsynligheten:

$$\hat{p} = \frac{X}{n} = \frac{\text{antall suksesser i løpet av } n \text{ forsøk}}{\text{antall forsøk}}$$

Så lenge n er stor nok, $n \geq 20$ (dersom X er hypergeometrisk må i tillegg n være liten nok i forhold til populasjonen $(N >> n)$), er $X \simeq N\left(np, \sqrt{np(1-p)}\right)$. Derfor blir $\hat{p}$ tilnærmet normalfordelt $N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$. Siden vi ikke kjenner verdien av p må vi bruke den estimerte verdien $\hat{p}$ når vi skal lage konfidensintervaller for p:

$$\hat{p} \pm (z\text{-verdi som er bestemt av konfidensnivået}) \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

3.4 Estimering av antall hendelser per tidsenhet λ

Hvis X er Poissonfordelt med forventningsverdi λ (ukjent) per tidsenhet, er

$$\hat{\lambda} = \frac{X}{t} = \frac{\text{antall hendeler i løpet av } t \text{ tidsenheter}}{\text{antall tidsenheter}}$$

Så lenge $\lambda t \geq 10$ er $X \simeq N(\lambda t, \sqrt{\lambda t})$ og dermed blir $\hat{\lambda} \simeq N\left(\lambda, \sqrt{\frac{\lambda}{t}}\right)$. Siden vi ikke kjenner verdien av λ, bruker vi $\hat{\lambda}$ når vi skal lage konfidensintervaller for λ:

$$\hat{\lambda} \pm (z\text{-verdi som er bestemt av konfidensnivået}) \cdot \sqrt{\frac{\hat{\lambda}}{t}}$$

4 Hypotesetesting på én dataserie

4.1 Z-test: Test av μ når σ er kjent

Testobservatoren er

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Du tror på forventningsverdien μ_0 inntil testen eventuelt viser at nullhypotesen skal forkastes. Kritisk z-verdi avhenger av konfidensnivået, og den finnes i tabellen for normalfordelingen:

	H_0	H_1	Forkast H_0 hvis
Alt. 1	$\mu \leq \mu_0$	$\mu > \mu_0$	$Z >$(kritisk z-verdi)
Alt. 2	$\mu \geq \mu_0$	$\mu < \mu_0$	$Z < -$(kritisk z-verdi)
Alt. 3	$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert Z\rvert >$(kritisk z-verdi)

4.2 T-test: Test av μ når σ er ukjent

Testobservatoren er

$$T = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

Du tror på forventningsverdien μ_0 inntil testen eventuelt viser at nullhypotesen skal forkastes. Kritisk t-verdi avhenger av konfidensnivået, og den finnes i tabellen for t-fordelingen med $(n-1)$ frihetsgrader:

	H_0	H_1	Forkast H_0 hvis
Alt. 1	$\mu \leq \mu_0$	$\mu > \mu_0$	$T >$(kritisk t-verdi)
Alt. 2	$\mu \geq \mu_0$	$\mu < \mu_0$	$T < -$(kritisk t-verdi)
Alt. 3	$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert T\rvert >$(kritisk t-verdi)

4.3 Hypotesetest av sannsynligheten p

Testobservatoren er

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}$$

Du tror på sannsynligheten p_0 inntil testen eventuelt viser at nullhypotesen skal forkastes. Kritisk z-verdi avhenger av konfidensnivået, og den finnes i tabellen for normalfordelingen:

	H_0	H_1	Forkast H_0 hvis
Alt. 1	$p \leq p_0$	$p > p_0$	$Z >$(kritisk z-verdi)
Alt. 2	$p \geq p_0$	$p < p_0$	$Z < -$(kritisk z-verdi)
Alt. 3	$p = p_0$	$p \neq p_0$	$\|Z\| >$(kritisk z-verdi)

4.4 Grubbs test for ensomme uteliggere

Hypoteser:

H_0: Det er ingen uteliggere i datasettet

H_1: Det er nøyaktig én uteligger i datasettet

Testobservatoren er

$$G = \frac{\max|Y_i - \overline{Y}|}{S}$$

der $Y_1, \ldots, Y_N$ er dataverdiene, $\overline{Y}$ er gjennomsnittet av dataverdiene og S er utvalgets standardavvik $\left(S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \overline{Y})^2\right)$. Nullhypotesen forkastes dersom

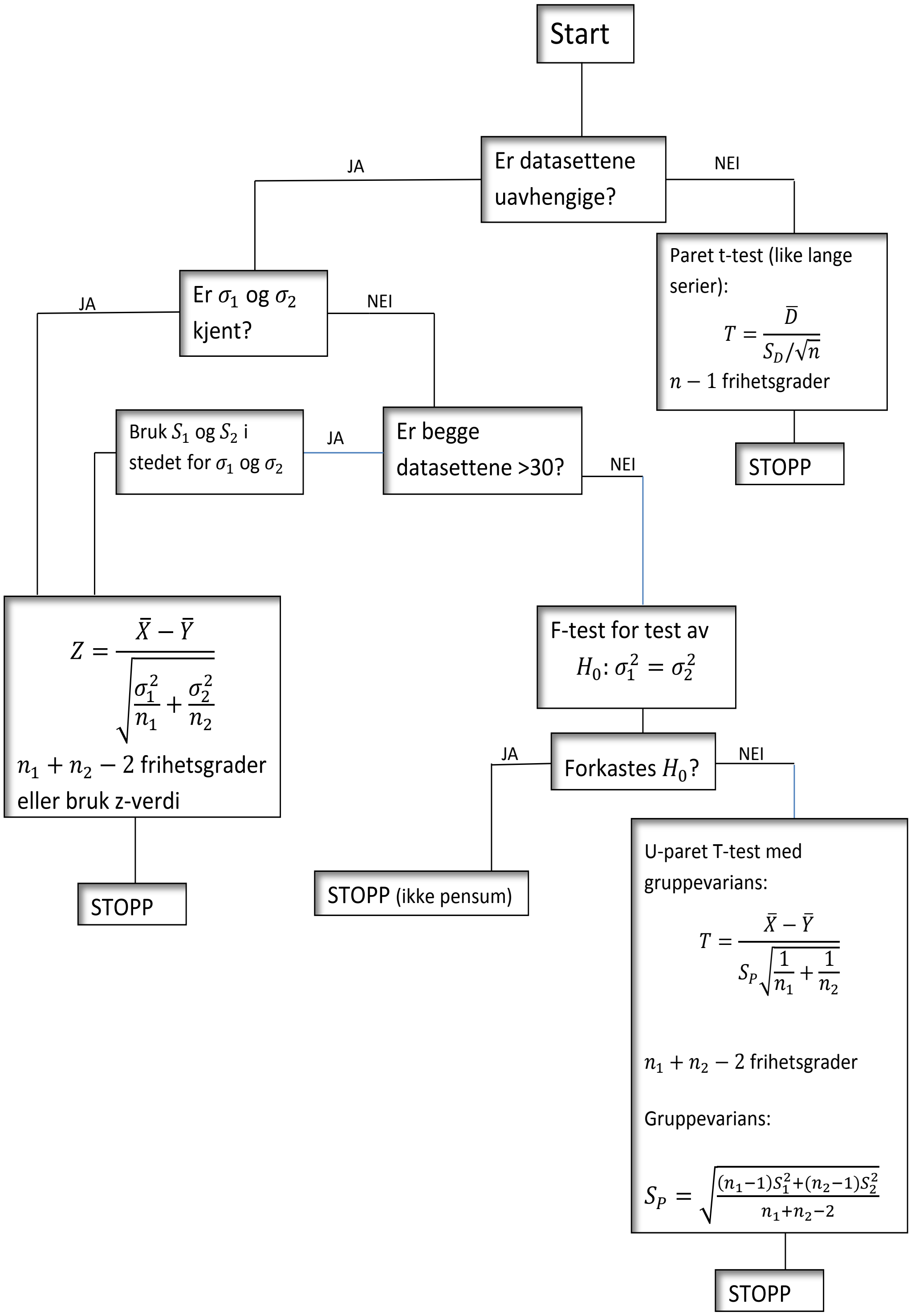
$$G > \frac{N-1}{\sqrt{N}}\sqrt{\frac{t^2}{N-2+t^2}}$$

der t finnes i tabellen for t-fordelingen:

* $N-2$ frihetsgrader
* signifikansnivå $\alpha/2N$

Ved ensidig test (sjekker om største/minste verdi er uteligger), brukes signifikansnivået α/N for å finne t.

Hypotesetesting med to dataserier



Kumulativ standardnormalfordeling

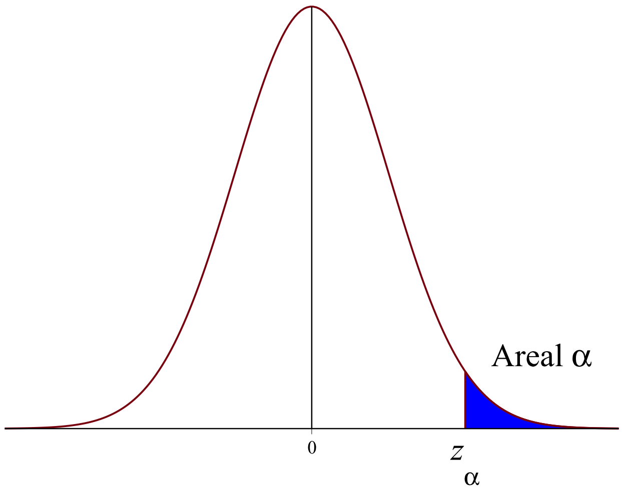
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,00	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,90	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,80	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,70	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,60	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,50	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,40	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,30	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,20	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,10	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,00	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,90	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,80	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,70	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,60	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,50	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,40	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,30	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,20	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,10	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,00	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,90	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,80	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,70	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,60	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,50	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,40	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,30	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,20	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,10	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,00	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,00	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,10	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,20	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,30	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,40	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,50	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,60	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,70	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,80	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,90	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,00	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,10	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,20	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,30	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,40	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,50	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,60	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,70	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,80	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,90	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,00	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,10	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,20	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,30	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,40	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,50	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,60	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,70	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,80	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,90	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,00	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

Elise Øby 2015

t-fordelingens kvantiltabell

Antall frihetsgrader	Areal *alfa* 0,25	0,1	0,05	0,025	0,01	0,005
1	1,000	3,078	6,314	12,706	31,821	63,657
2	0,816	1,886	2,920	4,303	6,965	9,925
3	0,765	1,638	2,353	3,182	4,541	5,841
4	0,741	1,533	2,132	2,776	3,747	4,604
5	0,727	1,476	2,015	2,571	3,365	4,032
6	0,718	1,440	1,943	2,447	3,143	3,707
7	0,711	1,415	1,895	2,365	2,998	3,499
8	0,706	1,397	1,860	2,306	2,896	3,355
9	0,703	1,383	1,833	2,262	2,821	3,250
10	0,700	1,372	1,812	2,228	2,764	3,169
11	0,697	1,363	1,796	2,201	2,718	3,106
12	0,695	1,356	1,782	2,179	2,681	3,055
13	0,694	1,350	1,771	2,160	2,650	3,012
14	0,692	1,345	1,761	2,145	2,624	2,977
15	0,691	1,341	1,753	2,131	2,602	2,947
16	0,690	1,337	1,746	2,120	2,583	2,921
17	0,689	1,333	1,740	2,110	2,567	2,898
18	0,688	1,330	1,734	2,101	2,552	2,878
19	0,688	1,328	1,729	2,093	2,539	2,861
20	0,687	1,325	1,725	2,086	2,528	2,845
21	0,686	1,323	1,721	2,080	2,518	2,831
22	0,686	1,321	1,717	2,074	2,508	2,819
23	0,685	1,319	1,714	2,069	2,500	2,807
24	0,685	1,318	1,711	2,064	2,492	2,797
25	0,684	1,316	1,708	2,060	2,485	2,787
26	0,684	1,315	1,706	2,056	2,479	2,779
27	0,684	1,314	1,703	2,052	2,473	2,771
28	0,683	1,313	1,701	2,048	2,467	2,763
29	0,683	1,311	1,699	2,045	2,462	2,756
30	0,683	1,310	1,697	2,042	2,457	2,750
31	0,682	1,309	1,696	2,040	2,453	2,744
32	0,682	1,309	1,694	2,037	2,449	2,738
33	0,682	1,308	1,692	2,035	2,445	2,733
34	0,682	1,307	1,691	2,032	2,441	2,728
35	0,682	1,306	1,690	2,030	2,438	2,724
40	0,681	1,303	1,684	2,021	2,423	2,704
45	0,680	1,301	1,679	2,014	2,412	2,690
50	0,679	1,299	1,676	2,009	2,403	2,678
60	0,679	1,296	1,671	2,000	2,390	2,660
70	0,678	1,294	1,667	1,994	2,381	2,648
80	0,678	1,292	1,664	1,990	2,374	2,639
100	0,677	1,290	1,660	1,984	2,364	2,626
1000	0,675	1,282	1,646	1,962	2,330	2,581
10000	0,675	1,282	1,645	1,960	2,327	2,576

Standardnormalfordelingens kvantiltabell



α	z_α
0,100	1,282
0,050	1,645
0,025	1,960
0,010	2,326
0,005	2,576
0,001	3,090