Estimation og test i normalfordelingen

Denne tekst indeholder et overblik over nogle grundlæggende principper for estimation og test i normalfordelingen i hyppigt forekommende situationer:

1 – én middelværdi, variansen kendt
2 – én middelværdi, variansen ukendt
3 – parvise observationer
4 – to middelværdier, samme varians
5 – to middelværdier, forskellig varians
6 – én varians (spredning)
7 – to varianser (spredninger)

Der er tale generelle statistiske teknikker, som nemt kan udføres ved hjælp af Microsoft Excel.

Formålet med disse teknikker er at udføre estimation af og test for middelværdi og varians i én eller to grupper af observationer for normalfordelte data. Alle disse teknikker er i øvrigt beskrevet i ISO 2854.

Hvis data ikke (tilnærmelsesvist) kan beskrives ved en normalfordeling, må man enten transformere data (f.eks. logaritmisk) eller anvende andre ("ikke-parametriske") statistiske teknikker. Er man i tvivl, om data kan beskrives tilfredsstillende ved hjælp af en normalfordeling, kan man evt. anvende grafisk tjek af normalitet (normalfordelingsplot). Dette behandles ikke i disse noter – der henvises til ISO 5479.

Det forudsættes, at læseren er i stand til at beregne gennemsnit og standardafvigelse for en gruppe af observationer. Dette kan f.eks. gøres i Excel ved hjælp af funktionerne AVERAGE henholdsvis STDEV. For god ordens skyld gengives formlerne her.

Af og til anvendes betegnelsen "middelværdi" (ofte betegnet μ) for en "sand" værdi, mens "gennemsnit" anvendes for den estimerede værdi. På tilsvarende måde skelnes af og til mellem begreberne "spredning" (ofte betegnet σ) for den sande værdi og "standardafvigelse" for den estimerede værdi. Ofte vil man dog anvende betegnelserne i flæng.

Vi har en gruppe af n uafhængige og normalfordelte observationer x_1 til x_n. Gennemsnittet ("middelværdien") er givet ved formlen

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Variansen er givet ved formlen

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

og standardafvigelsen ("spredningen") er *s*, dvs. kvadratroden af variansen.

I resten af disse noter vil der ikke blive givet detaljerede formler overalt.

I mange sammenhænge har man brug for mere avancerede statistiske metoder. Især er der ofte brug for "lineære modeller", herunder regressions- og variansanalyse. Dette er modeller, hvor man studerer indflydelsen af en eller flere faktorer på en afhængig variabel (interessevariablen). Disse faktorer kan være enten kvantitative (regressionsanalyse), kvalitative (variansanalyse) eller der kan evt. være begge typer på én gang ("generelle lineære modeller"). Disse modeller kan f.eks. analyseres ved hjælp af statistiksoftware som SAS/JMP.

1 – én middelværdi, variansen kendt

Baggrund

Vi har en gruppe af *n* uafhængige og normalfordelte observationer x_1 til x_n, jf. ovenfor. Deres *spredning* σ tænkes *kendt*, og det gælder dermed også variansen σ^2. Derimod er *middelværdien* μ *ukendt*.

Formålet er
- *At estimere middelværdien μ samt angive et konfidensinterval for denne.*
- *At teste, om middelværdien μ kan antages at være lig med en på forhånd given værdi μ_0.*

Vi skal derfor ikke foretage nogen beregning af standardafvigelsen.

Er man i tvivl, om den givne standardafvigelse stadig kan anvendes, kan man evt. teste dette. Se afsnit 6.

Eksempel

x_1	x_2	x_3	x_4		n	$\bar{x}$	μ_0	$\bar{x} - \mu_0$	σ	z_0
4,6	8,5	4,9	5,3		4	5,825	5	0,825	1,5	1,100

I dette eksempel er der 4 observationer, dvs. $n=4$. Vi ved fra mange tidligere analyser, at spredningen (standardafvigelsen) kan antages at være konstant $\sigma = 1{,}5$.

Estimation

Middelværdien (μ) er ukendt, men estimeres ved gennemsnittet $\bar{x} = 5{,}825$.

Konfidensinterval

Et gennemsnit er (ikke overraskende) mere sikkert bestemt, jo flere observationer man har. Hvis man f.eks. får 4 gange så mange observationer, skal variansen også divideres med 4, dvs. standardafvigelsen skal divideres med 2.

Generelt gælder der følgende regel: Standardafvigelsen for gennemsnittet får man ved at dividere den oprindelige standardafvigelse σ med kvadratroden af antallet af observationer n, dvs. den bliver $\sigma/\sqrt{n}$.

Vi vælger derfor et interval på formen $\mu \pm k \cdot \sigma/\sqrt{n}$, hvor konstanten k sættes til $z_{1-\alpha/2}$, dvs. en "fraktil" i (standard) normalfordelingen. Hvis f.eks. $\alpha = 0{,}05 = 5\%$, bliver
$z_{1-\alpha/2} = 1{,}96$. Dette hænger sammen med, at 95% af alle observationerne i en standard normalfordeling ligger mellem –1,96 og 1,96. Teknisk kalder man –1,96 for 2,5%-fraktilen og 1,96 for 97,5%-fraktilen i (standard) normalfordelingen.

Konfidensintervallet bliver derfor:

$$\bar{x} - 1{,}96 \cdot \sigma / \sqrt{n} \leq \mu \leq \bar{x} + 1{,}96 \cdot \sigma / \sqrt{n}$$

Dette interval kan beregnes til [4,36; 7,29], og det vil med 95% sandsynlighed indeholde den sande men ukendte værdi af middelværdien μ.

Konfidensintervallet kan generelt skrives:

$$\bar{x} - z_{1-\alpha/2}\sigma / \sqrt{n} \leq \mu \leq \bar{x} + z_{1-\alpha/2}\sigma / \sqrt{n}$$

Vil vi f.eks. i stedet have et interval, der med 99% sandsynlighed indeholder den ukendte værdi af middelværdien, skal vi i stedet vælge $\alpha = 0{,}01 = 1\%$, og så bliver
$z_{1-\alpha/2} = 2{,}576$, idet 99% af alle observationerne i en standard normalfordeling ligger mellem –2,576 (som er 0,5%-fraktilen) og 2,576 (som er 99,5%-fraktilen). Intervallet bliver i så fald [3,89; 7,76].

Beregning af teststørrelse

Vi skal nu teste, om gennemsnittet afviger fra en targetværdi på $\mu_0=5,0$.

Vores gennemsnit $\bar{x} = 5,825$ skal altså sammenlignes med targetværdien 5,0. Numerisk store (positive eller negative) afvigelser mellem disse størrelser fører til, at vi forkaster antagelsen om en middelværdi = 5,0.

Afvigelserne mellem gennemsnittet og target skal naturligvis ses i forhold til spredningen, dvs. $\sigma=1,5$. Hvis der er stor spredning, skal der (numerisk) større afvigelser til, før vi må forkaste antagelsen om en middelværdi = 5,0.

Som nævnt ovenfor fås standardafvigelsen af et gennemsnit ved at dividere den oprindelige standardafvigelse med kvadratroden af antallet af observationer.

Dette fører til, at man anvender følgende "teststørrelse":

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

I eksemplet ovenfor bliver z_0 = 1,100. Spørgsmålet er så, hvor stor en værdi (numerisk) af z_0 som fører til, at vi forkaster antagelsen om en middelværdi = 5,0.

Teststørrelsens fordeling

Dette spørgsmål hænger igen sammen med, hvilken fordeling vores teststørrelse følger. I det aktuelle tilfælde kan man vise, at teststørrelsen følger en "standard" normalfordeling (dvs. middelværdi 0 og spredning 1). Det har den fordel, at denne fordeling er tabelleret.

Man behøver derfor blot at slå op i en tabel over (standard) normalfordelingen for at se, hvor (numerisk) store værdier af z_0 som fører til forkastelse af antagelsen om en middelværdi = 5,0.

Svaret på dette spørgsmål hænger så igen sammen med valget af signifikansniveau. Oftest vælger vi signifikansniveau 5%, men af og til vælger vi signifikansniveau 1%. Det er disse to signifikansniveauer, som anbefales i forskellige ISO-standarder.

Hvis vi vælger et signifikansniveau på 5%, er (numeriske) værdier over 1,96 kritiske. Dette hænger sammen med, at 95% af alle observationerne i en standard normalfordeling ligger mellem –1,96 og 1,96. Teknisk kalder man –1,96 for 2,5%-fraktilen og 1,96 for 97,5%-fraktilen. Det er altså disse, man skal finde i tabellen.

Vælger vi et signifikansniveau på 1%, bliver (numeriske) værdier over 2,576 kritiske, idet 99% af observationerne i en standard normalfordeling ligger mellem –2,576 (som er 0,5%-fraktilen) og 2,576 (som er 99,5%-fraktilen).

Konklusion

Vi har i eksemplet fundet en teststørrelse på 1,100. Dette ligger pænt under de 1,96, hvorfor der ikke er statistisk belæg for at forkaste antagelsen om, at middelværdien er lig med targetværdien, denne antagelse må altså accepteres (på 5%-niveau, og naturligvis også på 1%-niveau).

På jævnt dansk: Den fundne gennemsnitsværdi 5,825 afviger ikke markant fra den accepterede targetværdi 5,0.

Ensidet test

I sjældne tilfælde er det sådan, at man på forhånd ved, at det er umuligt at få et gennemsnit under target. I så fald anvendes et såkaldt ensidet test.

I eksemplet gøres det sådan: Vi har et gennemsnit på 5,825, der jo som forventet er over de 5,0 (ellers var der noget galt med vores forhåndsviden!).

I et ensidet test skal man (for test på 5%-niveau) sammenligne teststørrelsen 1,100 med 95%-fraktilen, i stedet for 97,5%-fraktilen. I en tabel aflæses denne til 1,645. Vores antagelse om, at gennemsnittet er lig med target, bliver stadig accepteret.

For et test på 1%-niveau skal man sammenligne med 99%-fraktilen, som er 2,326.

Hvis vores forhåndsviden omvendt er, at det er umuligt at få gennemsnit over target, så er værdier under –1,645 (hhv. –2,326) kritiske.

2 – én middelværdi, variansen ukendt

Baggrund

Vi har som ovenfor en gruppe af n uafhængige og normalfordelte observationer x_1 til x_n. Deres *spredning* σ tænkes *ukendt*, det gælder dermed også variansen σ^2. Ligeledes er *middelværdien* μ *ukendt*.

Formålet er

– *At estimere middelværdien μ samt angive et konfidensinterval for denne.*
– *At teste, om middelværdien μ kan antages at være lig med en på forhånd given værdi μ_0.*

Eksempel

Samme data som i test nr. 1.

x_1	x_2	x_3	x_4		n	$\bar{x}$	μ_0	$\bar{x}$ - μ_0	s	t_0
4,6	8,5	4,9	5,3		4	5,825	5	0,825	1,806	0,914

I dette eksempel er der 4 observationer, dvs. n=4.

Estimation

Vi skal nu både estimere spredningen (standardafvigelsen) og middelværdien (gennemsnittet).

Middelværdien (μ) estimeres ved gennemsnittet $\bar{x}$ *= 5,825*.
Spredningen σ estimeres nu ved standardafvigelsen *s =1,806*.

Konfidensinterval

Vi skal her kun angive et konfidensinterval for middelværdien μ. I afsnit 6 vil vi angive et konfidensinterval for variansen hhv. standardafvigelsen.

I analogi med afsnit 1 opstiller vi et konfidensinterval for middelværdien μ af formen:

$$\bar{x} - t_{1-\alpha/2,n-1} s / \sqrt{n} \leq \mu \leq \bar{x} + t_{1-\alpha/2,n-1} s / \sqrt{n}$$

Forskellen er, at vi ikke kender spredningen σ, som i stedet estimeres ved standardafvigelsen *s.*

Det viser sig nu, at den konstant, man skal gange standardafvigelsen med, bliver noget større end før. I stedet for fraktiler i normalfordelingen skal man nu anvende en "Students" t-fordeling. Dette er ikke én, men en hel familie af fordelinger. Hvis der er *n* observationer (mindst 2), skal man anvende en (Students) t-fordeling med *n-1* "frihedsgrader".

I eksemplet er der 4 observationer, dvs. antal frihedsgrader er 3.

Ønsker vi et konfidensinterval, der med sandsynlighed 95% indeholder den sande men ukendte værdi af middelværdien μ, skal vi anvende 97,5%-fraktilen, som fås i en tabel til 3,182.

Intervallet beregnes derefter til [2,95; 8,70].

Ønsker vi et i stedet et konfidensinterval, der med sandsynlighed 99% indeholder den sande med ukendte værdi af middelværdien μ, skal vi anvende 99,5% fraktilen, som fås i en tabel til 5,841.

Intervallet beregnes derefter til [0,55; 11,10].

Disse konfidensintervaller er som man kan se temmelig meget bredere, end da vi kendte spredningen σ. Det er så at sige "straffen" for ikke at kende spredningen σ.

Beregning af teststørrelse

Vi skal nu teste, om dette gennemsnittet afviger fra en targetværdi på μ_0=5,0.

I forhold til test nr. 1 er det nye, at vi anvender standardafvigelsen s som mål for spredning, idet vi ikke har en på forhånd kendt værdi.

Dette fører til, at man anvender følgende teststørrelse, som vi nu kalder t_0:

$$t_0 = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

I eksemplet ovenfor bliver $t_0 = 0{,}914$. Spørgsmålet er så, hvor stor en værdi (numerisk) af t_0 som fører til, at vi forkaster antagelsen om en middelværdi = 5,0.

Teststørrelsens fordeling

Nu er der ikke længere tale om, at teststørrelsen følger en (standard) normalfordeling. Når antallet af observationer er stort (f.eks. over 20), vil det dog være en god tilnærmelse, men når der er tale om få observationer, gælder det ikke længere.

I stedet for normalfordelingen skal man nu anvende en "Students" t-fordeling med *n-1* "frihedsgrader", hvis der er *n* observationer.

I eksemplet er der 4 observationer, dvs. antal frihedsgrader er 3.

Hvis vi vælger et signifikansniveau på 5%, findes den kritiske grænse til 3,182 (97,5%-fraktilen) i tabellen.

Hvis vi vælger et signifikansniveau på 1%, findes den kritiske grænse til 5,841 (99,5%-fraktilen) i tabellen.

Konklusion

Vi har i eksemplet fundet en teststørrelse på 0,914. Dette ligger pænt under de 3,182, hvorfor der ikke er statistisk belæg for at forkaste antagelsen om, at middelværdien er lig med targetværdien. Denne antagelse må altså accepteres (på 5%-niveau, og naturligvis også på 1%-niveau).

Ensidet test

Hvis vi f.eks. på forhånd ved, at det er umuligt at få et gennemsnit under target, anvendes et ensidet test:

I eksemplet har vi et gennemsnit på 5,825, der jo som forventet er over de 5,0. I et ensidet test skal man (for test på 5%-niveau) sammenligne teststørrelsen 0,914 med 95%-fraktilen, i stedet for 97,5%-fraktilen. I tabel over t-fordelingen (med 3 frihedsgrader) aflæses denne til 2,353. Vores antagelse om, at gennemsnittet er lig med target, bliver stadig accepteret.

For et test på 1%-niveau skal man sammenligne med 99%-fraktilen, som er 4,541.

3 – parvise observationer

Baggrund

Situationen kan f.eks. være følgende: Vi har *n prøver*, som hver er blevet analyseret af *to laboratorier*. Vi er interesserede i at undersøge, *om der er forskelle mellem de to laboratorier (og hvor stor forskellen er)*, hvorimod forskellene mellem prøver i denne sammenhæng ikke er interessante.

Formålet er derfor

- *At estimere middelværdien μ_D for forskellen mellem laboratorier samt angive et konfidensinterval for denne.*
- *At teste, om middelværdien μ_D kan antages at være lig med 0.*

Dette problem kan håndteres ved at tage differensen mellem resultaterne for de to laboratorier, idet det kun er forskellen mellem laboratorierne, som er interessant. De enkelte resultater for hvert laboratorium er herefter ikke relevante.

Eksempel

Laboratorium 1	4,6	8,5	4,9	5,3				
Laboratorium 2	5,0	9,0	5,0	6,0	n	$\bar{d}$	s_d	t_0
Differens	0,4	0,5	0,1	0,7	4	0,4	0,25	3,4

Vi har 2 laboratorier, og der er foretaget analyser af 4 prøver. I første række er vist resultaterne for laboratorium 1, i anden række resultaterne for laboratorium 2.

Differensen mellem resultaterne (Laboratorium 2 – Laboratorium 1) er vist i 3. række, sammen med de relevante beregninger.

Estimation

Middelværdien μ_D for forskellen mellem laboratorier estimeres ved gennemsnittet af differenserne $\bar{d}$ = *0,4*.

Konfidensinterval

Et konfidensinterval for middelværdien μ_D af forskellen mellem laboratorier kan beregnes på samme måde som i afsnit 2, dvs.

$$\bar{d} - t_{1-\alpha/2,n-1} s_d / \sqrt{n} \le \mu_d \le \bar{d} + t_{1-\alpha/2,n-1} s_d / \sqrt{n}$$

I eksemplet er der 4 differenser, dvs. antal frihedsgrader er 3.

Ønsker vi et konfidensinterval, der med sandsynlighed 95% indeholder den sande men ukendte værdi af middelværdien μ, skal vi anvende 97,5%-fraktilen 3,182.

Intervallet beregnes derefter til [0,03; 0,82].

Ønsker vi i stedet et konfidensinterval, der med sandsynlighed 99% indeholder den sande med ukendte værdi af middelværdien μ, skal vi anvende 99,5%-fraktilen, som fås i en tabel til 5,841.

Intervallet beregnes derefter til [-0,31; 1,16].

Beregning af teststørrelse

Testet er magen til test nr. 2, idet spredningen af differenserne i praksis vil være ukendt. Vi kan derfor opstille teststørrelsen

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

I eksemplet får vi $t_0 = 3{,}4$.

Teststørrelsens fordeling

Vi skal teste, om laboratorierne er ens, dvs. om middelværdien af differenserne kan antages at være 0. I forhold til beskrivelsen af test nr. 2 svarer 0 derfor til target middelværdien.

Som for test nr. 2 følger denne teststørrelse en (Students) t-fordeling. Antallet af frihedsgrader er også her n-1, idet vi har n differenser.

I eksemplet er der 4 differenser, dvs. antal frihedsgrader er 3.

Kritiske værdier er (som i eksemplet for test nr. 2) hhv. 3,182 (97,5%-fraktilen, signifikansniveau på 5%) og 5,841 (99,5%-fraktilen, signifikansniveau på 1%).

Konklusion

Vi har i eksemplet fundet en teststørrelse på 3,4. Dette ligger over 3,182, hvorfor vi må forkaste antagelsen om, at laboratorierne er ens (på 5%-niveau).

Derimod ligger teststørrelsen 3,4 under 5,841. På 1%-signifikansniveau vil vi altså acceptere antagelsen om, at laboratorierne er ens.

Ensidet test

Hvis vi f.eks. på forhånd ved (af faglige grunde), at laboratorium 2 altid vil give resultater større end resultaterne for laboratorium 1, anvendes et ensidet test:

Kritiske værdier fås (som i eksemplet for test nr. 2) til hhv. 2,353 (5%-niveau, 95%-fraktilen) og 4,541 (1%-niveau, 99%-fraktilen).

Konklusionen bliver her den samme som for et tosidet test: På 5%-niveau forkaster vi, mens vi på 1%-niveau accepterer antagelsen om, at laboratorierne er ens.

4 – to middelværdier, samme varians

Baggrund

Situationen er nu følgende: Vi har to grupper af uafhængige og normalfordelte observationer. *Deres spredning σ tænkes ukendt, men fælles* for begge grupper. Middelværdien i begge grupper er ukendt.

Formålet er

- *At estimere forskellen mellem middelværdien i de to grupper samt angive et konfidensinterval for denne.*
- *At teste, om middelværdien i de to grupper kan antages identiske.*

Er man i tvivl, om det er rimeligt at antage, at de to grupper har samme spredning, kan man evt. anvende testet i afsnit 7.

Eksempel

Vi har f.eks. 4 gentagne målinger foretaget på én prøve og 3 gentagne målinger foretaget på en anden prøve. Vi vil undersøge, om de to prøver kan antages at være identiske.

						$\bar{x}_i$	n_i	s_i
Prøve 1	3,6	4,5	3,9	4,3		4,075	4	0,403
Prøve 2	4,6	5,2	4,9			4,900	3	0,300

Her er anført gennemsnit, antal målinger samt standardafvigelse for hver prøve.

Bemærk:

- Antallet af observationer i de to grupper behøver ikke at være identiske!
- To målinger i samme kolonne har intet med hinanden at gøre! Målingerne fra samme prøve kan byttes rundt frit, de står i vilkårlig rækkefølge.

Estimation

I dette afsnit skal vi estimere middelværdien for begge grupper hver for sig, differensen mellem disse samt den fælles spredning.

Middelværdien i gruppe 1 (μ_1) estimeres ved $\bar{x}_1$ (gennemsnittet for prøve nr. 1), tilsvarende estimeres middelværdien i gruppe 2 (μ_2) ved $\bar{x}_2$ (gennemsnittet for prøve nr. 2).

Differensen μ_1-μ_2 estimeres ved $\bar{x}_1 - \bar{x}_2$ *= -0,825.*

Den fælles spredning estimeres som en "poolet" ("gennemsnitlig") standardafvigelse for de to prøver ved hjælp af følgende formel:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Her er s_1 hhv. n_1 standardafvigelse og antal målinger i første prøve, og tilsvarende haves s_2 henh. n_2 for prøve nr. 2. I eksemplet fås $s_p = 0{,}365$.

Konfidensinterval

Et konfidensinterval for forskellen mellem middelværdier, dvs. differensen μ_1-μ_2, kan beregnes ved hjælp af følgende formel:

$$\bar{x}_1 - \bar{x}_2 - t_{1-\alpha/2, n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2$$
$$\le \bar{x}_1 - \bar{x}_2 + t_{1-\alpha/2, n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Her er konstanten $t_{1-\alpha/2, n_1+n_2-2}$ den relevante fraktil i en t-fordeling med $n_1 + n_2 - 2$ frihedsgrader, dvs. i eksemplet er der 5 frihedsgrader.

Denne fraktil kan aflæses i tabel til 2,571 (for et 95%-konfidensinterval) hhv. 4,032 (for et 99%-konfidensinterval).

I eksemplet får vi et 95%-konfidensinterval [-1,54; -0,11] hhv. et 99%-konfidensinterval [-1,95; 0,30].

Beregning af teststørrelse

Vi skal nu teste, om forskellen mellem middelværdier μ_1 - μ_2 kan antages at være 0.

Først beregnes en teststørrelse

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Her er $\bar{x}_1$ gennemsnittet for prøve nr. 1, $\bar{x}_2$ er gennemsnittet for prøve nr. 2, mens s_p er fundet ovenfor. I eksemplet fås $t_0 = -2{,}959$.

Teststørrelsens fordeling

Igen følger teststørrelsen en t-fordeling. Antallet af frihedsgrader er $n_1 + n_2 - 2$, dvs. i eksemplet er der 5 frihedsgrader.

Kritiske værdier er hhv. 2,571 (signifikansniveau 5%) og 4,032 (signifikansniveau 1%).

Konklusion

Vi har i eksemplet fundet en teststørrelse på –2,959. Dette ligger uden for intervallet fra -2,571 til 2,571, hvorfor vi forkaster antagelsen om, at middelværdierne er identiske (på 5%-niveau).

Derimod ligger teststørrelsen –2,959 inden for intervallet –4,032 til 4,032, hvorfor vi på 1%-niveau må acceptere antagelsen om, at middelværdierne er identiske.

Ensidet test

Hvis vi på forhånd ved (af faglige grunde), at prøve nr. 2 altid vil give større måleværdier end prøve nr. 1, anvendes et ensidet test:

Kritiske værdier fås til hhv. ±2,015 (5% niveau) og ±3,365 (1%-niveau). Konklusionen bliver her den samme som for et tosidet test.

5 – to middelværdier, forskellig varians

Baggrund

Her er situationen som i afsnit 4, men vi kan (eller vil) ikke antage, at variansen (spredningen) i de to grupper er identiske.

Formålet er

- *At estimere middelværdien for forskellen mellem grupper samt angive et konfidensinterval for denne.*
- *At teste, om middelværdien i de to grupper kan antages identiske.*

I mange situationer vil spredningen afhænge af middelværdien ("niveauet"). Hvis man på forhånd ikke ved, om middelværdierne er ens, kan man ikke automatisk antage, at spredningerne er ens.

Det kan også være, at man først har undersøgt, om spredningerne er ens (vha. testet i afsnit 7), men dette er blevet forkastet.

Eksempel

Samme eksempel som i afsnit 4. Blot vil vi ikke antage, at de to prøver har samme spredning.

						$\bar{x}_i$	n_i	s_i
Prøve 1	3,6	4,5	3,9	4,3		4,075	4	0,403
Prøve 2	4,6	5,2	4,9			4,900	3	0,300

Estimation

I dette afsnit skal vi estimere middelværdien og spredning for begge grupper hver for sig, samt endvidere differensen mellem middelværdierne.

Middelværdien i gruppe 1 (μ_1) estimeres ved $\bar{x}_1$ (gennemsnittet for prøve nr. 1), tilsvarende estimeres middelværdien i gruppe 2 (μ_2) ved $\bar{x}_2$ (gennemsnittet for prøve nr. 2).

Differensen μ_1-μ_2 estimeres ved $\bar{x}_1 - \bar{x}_2$ *= -0,825.*

Spredningen i gruppe 1 (σ_1) estimeres ved standardafvigelsen $s_1=0,403$, spredningen i gruppe 2 (σ_2) estimeres ved standardafvigelsen $s_2=0,300$.

Konfidensinterval

Et konfidensinterval for forskellen mellem middelværdier, dvs. differensen μ_1-μ_2, kan beregnes ved hjælp af følgende formel:

$$\bar{x}_1 - \bar{x}_2 - t_{1-\alpha/2,\upsilon}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \le \mu_1 - \mu_2$$

$$\le \bar{x}_1 - \bar{x}_2 + t_{1-\alpha/2,\upsilon}\sqrt{\frac{s_1^2}{n_1}+\frac{s_1^2}{n_2}}$$

Her er konstanten $t_{1-\alpha/2,\upsilon}$ den relevante fraktil i en t-fordeling med υ frihedsgrader.

Antallet af frihedsgrader er mere kompliceret end før:

$$\upsilon = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

Her får vi i eksemplet *υ = 4,987*. Dette afrundes til υ = 5, hvilket er samme antal frihedsgrader som i afsnit 4. Derfor bliver t-fraktilerne de samme som i afsnit 4, dvs. *2,571* (for et 95%-konfidensinterval) hhv. *4,032* (for et 99%-konfidensinterval).

I eksemplet får vi et 95%-konfidensinterval [-1,51; -0,14] hhv. et 99%-konfidensinterval [-1,90; 0,25].

Beregning af teststørrelse

Nu skal vi ikke beregne en fælles standardafvigelse, men beregner direkte følgende teststørrelse:

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

I teststørrelsen indgår altså udelukkende gennemsnit, standardafvigelse og antal målinger for hver prøve. I eksemplet får vi *t_0 = -3,10*.

Teststørrelsens fordeling

Her følger teststørrelsen med tilnærmelse en t-fordeling. Antallet af frihedsgrader (υ) er bestemt ovenfor til 5.

Kritiske værdier er hhv. ±2,571 (signifikansniveau på 5%) og ±4,032 (signifikansniveau på 1%).

Konklusion

Vi har i eksemplet fundet en teststørrelse på -3,10. Dette ligger uden for intervallet fra -2,571 til 2,571, hvorfor vi forkaster antagelsen om, at middelværdierne er identiske (på 5%-niveau).

Derimod ligger teststørrelsen -3,10 inden for intervallet –4,032 til 4,032, hvorfor vi på 1%-niveau må acceptere antagelsen om, at middelværdierne er identiske.

Ensidet test

Hvis vi på forhånd ved (af faglige grunde), at prøve nr. 2 altid vil give større måleværdier end prøve nr. 1, anvendes et ensidet test:

Kritiske værdier fås til hhv. ±2,015 (5%-niveau) og ±3,365 (1%-niveau).

Konklusionen bliver her den samme som for et tosidet test.

6 – én varians (spredning)

Baggrund

Vi har en gruppe af *n* uafhængige og normalfordelte observationer x_1 til x_n. Deres *spredning* σ tænkes *ukendt*, det gælder dermed også variansen σ^2. Ligeledes er *middelværdien* μ *ukendt*.

Formålet er

- *At estimere variansen σ^2 (eller spredningen σ) samt angive et konfidensinterval for denne.*
- *At teste, om variansen σ^2 kan antages at være lig med en på forhånd given værdi σ_0^2 (eller om $\sigma=\sigma_0$).*

Dette test kan f.eks. anvendes forud for test nr. 1, hvis man ikke er sikker på, om den kendte værdi af spredningen stadig kan anvendes.

Eksempel

Vi anvender samme data som for test nr. 1 og 2.

x_1	x_2	x_3	x_4		n	$\bar{x}$	s	σ_0	χ_0^2
4,6	8,5	4,9	5,3		4	5,825	1,806	1,5	4,35

I dette eksempel er der 4 observationer, dvs. n=4.

Estimation

Vi skal først estimere både spredningen (standardafvigelsen) og middelværdien (gennemsnittet).

Middelværdien (μ) estimeres ved gennemsnittet $\bar{x}$ *= 5,825*.
Spredningen σ estimeres nu ved standardafvigelsen *s =1,806*.

Konfidensinterval

Først konfidensinterval for variansen:

I dette tilfælde er konfidensintervallet specificeret direkte ved sin nedre hhv. øvre grænse (dvs. det er **ikke** et symmetrisk interval omkring variansen).

Vi får nu brug for en såkaldt χ^2-fordeling (Chi-i-anden, "chi-square"). Ligesom t-fordelingen er dette en familie af fordelinger, hver med sit antal frihedsgrader. Her er antal frihedsgrader *(n-1)*.

I eksemplet er n=4, dvs. antal frihedsgrader er 3.

Konfidensintervallet for variansen er:

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}$$

For et 95% konfidensinterval er nævneren i brøken hhv. 97,5%-fraktilen (nedre grænse) og 2,5%-fraktilen (øvre grænse) i en χ^2-fordeling med 3 frihedsgrader. Disse kan aflæses i tabel til hhv. 0,22 og 9,35.

Konfidensintervallet fås da til [1,05; 45,36].

For et 99%-konfidensinterval er nævneren i brøken hhv. 99,5%-fraktilen (nedre grænse) og 0,5%-fraktilen (øvre grænse) i en χ^2-fordeling med 3 frihedsgrader. Disse kan aflæses i tabel til hhv. 0,072 og 12,84.

Konfidensintervallet fås da til [0,76; 136,46].

Konfidensinterval for spredningen fås simpelthen ved at tage kvadratroden af grænserne i konfidensinterval for variansen.

Dermed får vi følgende konfidensintervaller for spredningen:

95% konfidensinterval: [1,02; 6,73].

99% konfidensinterval: [0,87;11,68].

Beregning af teststørrelse

Vi antager fra mange tidligere analyser, at spredningen (standardafvigelsen) er konstant =1,5 (dvs. om σ=1,5). Vores formål er at teste, om denne antagelse stadig holder.

Det vil være nærliggende at anvende forholdet s^2/σ_0^2 som teststørrelse. Her vil værdier langt fra 1 være kritiske. Af tekniske årsager vil man gange denne teststørrelse med *(n-1)*, hvor n = antal observationer.

Dvs. man anvender teststørrelsen

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Her vil værdier langt fra *(n-1)* være kritiske.

I eksemplet får vi χ^2_0=4,35.

Teststørrelsens fordeling

Teststørrelsen følger χ^2-fordeling med *n-1 = 3* frihedsgrader.

Både små og store værdier af χ_0^2 er kritiske. På 5%-signifikansniveau kan de kritiske værdier aflæses til 0,22 (2,5%-fraktilen) og 9,35 (97,5%-fraktilen). På 1%-signifikansniveau fås tilsvarende de kritiske værdier 0,07 (0,5%-fraktilen) og 12,84 (99,5%-fraktilen).

Konklusion

I eksemplet ligger teststørrelsen χ_0^2 = 4,35 mellem de kritiske værdier på 5%-signifikansniveau 0,22 og 9,35. Dette betyder, at vi vil acceptere antagelsen om, at spredningen er 1,5 på 5%-niveau (og dermed også på 1%-niveau).

Ensidet test

Hvis man på forhånd ved, at standardafvigelsen ikke kan være mindre (hhv. større) end den givne værdi, kan man anvende et ensidet test.

Her bliver den kritiske værdi på 5%-signifikansniveau 95%-fraktilen (hhv. 5%-fraktilen). På 1%-signifikansniveau skal man anvende 99%-fraktilen (hhv. 1%-fraktilen).

I eksemplet er den beregnede standardafvigelse s=1,806 større end den på forhånd givne værdi σ_0=1,5. Hvis dette er noget, man vidste på forhånd, kan man anvende et ensidet test, hvor man skal sammenligne teststørrelsen χ_0^2=4,35 med 95%-fraktilen i en χ^2-fordeling med 3 frihedsgrader. Denne kan aflæses i tabel til 7,81. I dette tilfælde vil vi derfor acceptere antagelsen om, at spredningen er lig med den givne værdi 1,5.

7 – to varianser (spredninger)

Vi har to grupper af uafhængige og normalfordelte observationer.

Formålet er

- *At estimere forholdet mellem variansen i de to grupper samt angive et konfidensinterval for dette.*
- *At teste, om variansen i de to grupper kan antages identiske.*

Dette test kan man f.eks. anvende forud for test nr. 4 hvis man ikke er sikker på, om de to grupper har samme varians.

Eksempel

Vi anvender samme data som i afsnit 4. Vi vil undersøge, om variansen (spredningen) i de to prøver kan antages at være identiske.

						$\bar{x}_i$	n_i	s_i
Prøve 1	3,6	4,5	3,9	4,3		4,075	4	0,403
Prøve 2	4,6	5,2	4,9			4,900	3	0,300

Her er anført gennemsnit, antal målinger samt standardafvigelse for hver prøve.

Estimation

I dette afsnit skal vi estimere middelværdien og spredning for begge grupper hver for sig, samt endvidere forholdet mellem varianserne.

Middelværdien i gruppe 1 (μ_1) estimeres ved $\bar{x}_1$ (gennemsnittet for prøve nr. 1), tilsvarende estimeres middelværdien i gruppe 2 (μ_2) ved $\bar{x}_2$ (gennemsnittet for prøve nr. 2).

Spredningen i gruppe 1 (σ_1) estimeres ved standardafvigelsen $s_1=0,403$, spredningen i gruppe 2 (σ_2) estimeres ved standardafvigelsen $s_2=0,300$.

Dvs. at forholdet mellem varianserne estimeres ved

$$\frac{s_1^2}{s_2^2}$$

Denne størrelse beregnes til 1,806.

Konfidensinterval

Vi får nu brug for en "F-fordeling". Dette er en familie af fordelinger, som har to antal frihedsgrader, et for tælleren og et for nævneren.

Dvs. vi skal anvende en F-fordeling med (3,2) frihedsgrader, idet antal frihedsgrader for tælleren s_1^2 er 3 (vi har 4 observationer), mens antal frihedsgrader for nævneren s_2^2 er 2 (vi har 3 observationer).

Konfidensintervallet for $\frac{\sigma_1^2}{\sigma_2^2}$ bliver

$$\frac{s_1^2}{s_2^2} \cdot F_{\alpha/2,n_1-1,n_2-1} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} F_{1-\alpha/2,n_1-1,n_2-1}$$

For et 95%-konfidensinterval anvender vi 97,5%-fraktilen i en F-fordeling med (3,2) frihedsgrader. Denne kan aflæses i tabel til 39,2.
Endvidere har vi brug for 2,5%-fraktilen i en F-fordeling med (3,2) frihedsgrader. Denne kan fås som det reciprokke af 97,5%-fraktilen i en F-fordeling med (2,3) frihedsgrader (NB: Byt om på frihedsgraderne!). Man får derved 2,5%-fraktilen til 1/16,0=0,0625.

Konfidensintervallet fås da til [0,11; 70,8].

For et 99%-konfidensinterval anvender vi 99,5%-fraktilen i en F-fordeling med (3,2) frihedsgrader. Denne kan aflæses i tabel til 199.
Endvidere har vi brug for 0,5%-fraktilen i en F-fordeling med (3,2) frihedsgrader. Denne kan fås som det reciprokke af 99,5%-fraktilen i en F-fordeling med (2,3) frihedsgrader (NB: Byt om på frihedsgraderne!). Man får derved 2,5%-fraktilen til 1/49,8=0,0201.

Konfidensintervallet fås da til [0,036;359].

Beregning af teststørrelse

Det vil være nærliggende at anvende forholdet mellem de to beregnede varianser som teststørrelse.

Dvs. vi anvender teststørrelsen

$$F_0 = s_1^2 / s_2^2$$

Værdier langt fra 1 er kritiske.

Man kunne lige så godt anvende teststørrelsen $1/F_0 = s_2^2 / s_1^2$. Af tekniske årsager vil man sørge for altid at anvende den største af disse to størrelser.

I eksemplet får vi F_0 =1,805. Dermed er $1/F_0$ = 0,554. I dette tilfælde anvendes F_0.

Teststørrelsens fordeling

Teststørrelsen følger en "F-fordeling". Dette er en familie af fordelinger, som har to antal frihedsgrader, et for tælleren og et for nævneren.

I eksemplet anvender vi $F_0 = s_1^2 / s_2^2$. Dvs. vi skal anvende en F-fordeling med (3,2) frihedsgrader, idet antal frihedsgrader for tælleren s_1^2 er 3 (vi har 4 observationer), mens antal frihedsgrader for nævneren s_2^2 er 2 (vi har 3 observationer).

Da vi har valgt den største af de to mulige teststørrelser F_0 og $1/F_0$, er det kun store værdier, som er kritiske. Men der er stadig tale om et tosidet test. Derfor skal vi kun sammenligne teststørrelsen med 97,5%-fraktilen i F-fordelingen (5%-signifikansniveau) *= 39,17* hhv. 99,5%-fraktilen (1%-signifikansniveau) *= 199,16*.

Konklusion

I eksemplet er teststørrelsen F_0 =1,805 mindre end 39,17 (97,5%-fraktilen), vil vi acceptere antagelsen om, at de to varianser (spredninger) er identiske.

Ensidet test

Antag f.eks., at vi på forhånd vidste, at variansen (spredningen) i gruppe 1 ville være større end i gruppe 2. I så fald skal vi i stedet sammenligne med 95%-fraktilen (test på 5%-signifikansniveau) hhv. 99%-fraktilen (1%-niveau).

I tabel over F-fordelingen med (3,2) frihedsgrader aflæses 95%-fraktilen til 19,16, mens 99%-fraktilen aflæses til 99,16. Konklusionen bliver stadig, at vi accepterer antagelsen om, at de to varianser (spredninger) er identiske.

Referencer

ISO 2854: *Statistical Interpretation of data – Techniques of estimation and tests relating to means and variances*. (Omhandler alle test, nr. 1 – nr. 7).

ISO 3301: *Statistical Interpretation of data – Comparison of two means in the case of paired observations*. (Omhandler specielt test nr. 3 mere detaljeret).