HD Dataanalyse, efterår 2018 Opgaveark 7



Baggrund og databeskrivelse (opgave 1-2)

Denne opgave ser nærmere på forskellige regnskabsnøgletal for et udsnit af små, mellem-store og store danske virksomheder. Opgaven tager udgangspunkt i filen Experian2008.jmp, hvis indhold er beskrevet på opgaveark 5.

Opgave 1 (egenkapitalens forrentning)

Opgave 1 omhandler udelukkende virksomheder med en egenkapitalforrentning på mellem -100% og 100%.

JMP-vink: “Rows” -> “Data Filter”

a). Tegn et histogram over fordelingen af egenkapitalens forrentning (Egenkapitalens forrentning (i %)) for henholdsvis anparts- og aktieselskaber (Selskabsform).

JMP-vink: “Analyze” -> “Distribution”. Placér Selskabsform under “By”.

Estimér middelværdi og standardafvigelse af virksomhedernes egenkapitalforrent-ning for henholdsvis anparts- og aktieselskaber.

JMP-vink: “Analyze” -> “Distribution”. Placér Selskabsform under “By”.



Normal(12.7421,36.8814) A/S

Normal(11.6639,39.2062) ApS

b). Beregn et 95%-konfidensinterval for forskellen i den forventede egenkapitalforrentning mellem anparts- og aktieselskaber og giv en fortolkning af intervallet.

JMP-vink: “Analyze” -> “Fit Y by X”. Placér Egenkapitalens forrentning (i %) under “Y, Response”,

Selskabsform under “X, Factor”. Vælg “t Test” via den røde trekant.



Vi kan med 95% sikkerhed sige at forskellen i egenkapitalforrentningen mellem ApS – A/S i populationen ligger mellem -2.5823% og 0.4259%. Bemærk 0 er indeholdt i intervallet for forskellen, vi kan altså ikke afvise at forskellen er 0 (dvs. der er ikke forskel i egenkapitalforrentningen mellem ApS – A/S.

c). Gør rede for om der er statistisk belæg for en nulhypotese om, at den forventede egenkapitalforrentning er den samme blandt anparts- og aktieselskaber (brug = 5% som signifikansniveau). Angiv i den forbindelse antagelser, hypoteser, signifikans-sandsynlighed1samt teknisk og let forståelig konklusion.

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Means”. Vælg “Raw

Data”.

|  |  |
| --- | --- |
| Test 2 Middelværdier ens varians, tosidet alternativ hypotese | |
| Forudsætninger | Vi gennemfører en Unpooled t-test. |
| Hypotese |  |
| P-værdi/signifikanssandsynlighed | 16% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/ signifikanssandsynligheden 16% er større end 5% signifikansniveauet forkaster vi ikke nulhypotesen. |
| Ikke-teknisk konklusion | Den forventede egenkapitalforrentning er altså den samme blandt anparts- og aktieselskaber |
| Det hurtigste er at bruge output fra forrige opgave hvor der er p-værdier enten pooled (ens varianser) eller unpooled t-test (uens varianser) Brug kun et test, alt er vist her kun for at demonstrere mulighederne :O)    Hvis vi benytter “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Means”. Vælg “Raw Data” får vi flere muligheder for at teste fx hvor stor forskellen er osv. Bemærk vi får samme p-værdi: | |

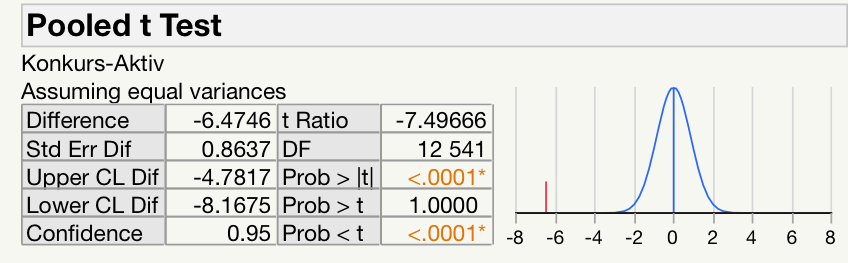
Gør rede for sammenhængen med resultatet i delspg. b).

Vi ser af KI for forskellen i B at vi ikke kan afvise ens egenkapitalforrentning, hvilket vi også konkluderer i c

d). Beregn et 95%-konfidensinterval for forskellen i den forventede egenkapitalforrentning mellem aktive og konkursramte selskaber (Status) og giv en fortolkning af intervallet.

JMP-vink: “Analyze” -> “Fit Y by X”. Placér Egenkapitalens forrentning (i %) under “Y, Response”,

Status under “X, Factor”. Vælg “t Test” via den røde trekant.



Vi kan med 95% sikkerhed sige at forskellen i egenkapitalforrentningen mellem Konkurs – Aktiv i populationen ligger mellem -8.33% og -4.62% (forskellen er negativ). Bemærk 0 er ikke indeholdt i intervallet for forskellen, vi kan altså afvise at forskellen er 0 (dvs. der er forskel i egenkapitalforrentningen mellem Konkurs – Aktiv.



1/6

HD Dataanalyse, efterår 2018 Opgaveark 7



e). Gør rede for om der er statistisk belæg for en nulhypotese om, at den forventede egenkapitalforrentning er den samme blandt aktive og konkursramte selskaber (brug

= 5% som signifikansniveau). Angiv i den forbindelse antagelser, hypoteser, signifi-kanssandsynlighed samt teknisk og let forståelig konklusion.

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Means”. Vælg “Raw

Data”.

|  |  |
| --- | --- |
| Test 2 Middelværdier ens varians, tosidet alternativ hypotese | |
| Forudsætninger | Det forudsættes at varianserne er ens for de 2 populationer. Vi gennemfører en pooled t-test  Forudsætning for test for varianshomogenitet er at begge populationer er normalfordelte. |
| Hypotese |  |
| Teststørrelse | -6.84 (unequal variances) |
| P-værdi/signifikanssandsynlighed | <0.0001 |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/ signifikanssandsynligheden <0.0001 er mindre end 5% signifikansniveauet afviser vi nulhypotesen. |
| Ikke-teknisk konklusion | Den forventede egenkapitalforrentning er ikke den samme blandt aktive og konkursramte selskaber. |
| Ved brug af calculator | |

Gør rede for om der er statistisk belæg for en nulhypotese om, at den forventede egenkapitalforrentning er 4% højere for aktive end for konkursramte selskaber (brug

= 5% som signifikansniveau).

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Means” Vælg. “Raw Data”.

|  |  |
| --- | --- |
| Test 2 Middelværdier ens varians, tosidet alternativ hypotese | |
| Forudsætninger | Det forudsættes at varianserne er ens for de 2 populationer. Vi gennemfører en pooled t-test  Forudsætning for test for varianshomogenitet er at begge populationer er normalfordelte. |
| Hypotese |  |
| P-værdi/signifikanssandsynlighed | 0.89% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/ signifikanssandsynligheden 0.89% er mindre end 5% signifikansniveauet afviser vi nulhypotesen. |
| Ikke-teknisk konklusion | Den forventede egenkapitalforrentning er altså ikke 4 procentpoint højere for aktive end for konkursramte selskaber |
| Ved brug af calculator: | |

Gør rede for om der er statistisk belæg for en nulhypotese om, at den forventede egenkapitalforrentning er 6% højere for aktive end for konkursramte selskaber (brug

= 5% som signifikansniveau).

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Means”. Vælg “Raw

Data”.

|  |  |
| --- | --- |
| Test 2 Middelværdier ens varians, tosidet alternativ hypotese | |
| Forudsætninger | Det forudsættes at varianserne er ens for de 2 populationer. Vi gennemfører en pooled t-test  Forudsætning for test for varianshomogenitet er at begge populationer er normalfordelte. |
| Hypotese |  |
| P-værdi/signifikanssandsynlighed | 61.58% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/ signifikanssandsynligheden 61.58% er større end 5% signifikansniveauet afviser vi ikke nulhypotesen. |
| Ikke-teknisk konklusion | Der er statistisk belæg for en nulhypotese om at egenkapitalforrentningen er 6 procentpoint højere for aktive end for konkursramte selskaber. |
| Ved brug af calculator: | |

Gør rede for sammenhængen mellem de ovenstående hypotesetest og konfidensin-tervallet i delspg. d).

Vi kunne i d) med 95% sikkerhed sige at forskellen i egenkapitalforrentningen mellem Konkurs – Aktiv i populationen ligger mellem -8.33% og -4.62%. -4% var ikke indeholdt i KI det er derimod -6%, hvorfor vi tror på at Aktive selskaber har en gennemsnitlig egenkapitalforrentning der er 6% højere end blandt selskaber under konkursbehandling.

* “Signifikanssandsynlighed” og “P-værdi” er identiske betegnelser.



2/6

HD Dataanalyse, efterår 2018 Opgaveark 7



Opgave 2 (dækningsgrad)

Opgave 2 omhandler udelukkende virksomheder med en egenkapitalforrentning på mellem -100% og 100%.

JMP-vink: “Rows” -> “Data Filter”

a). Tegn et box plot over fordelingen af dækningsgraden (Dækningsgrad (i %)) for hen-holdsvis anparts- og aktieselskaber (Selskabsform).

JMP-vink: “Analyze” -> “Fit Y by X”. Placér Dækningsgrad (i %) under “Y, Response”, Selskabsform

under “X, Factor”.



Estimér middelværdi og standardafvigelse af virksomhedernes dækningsgrad for henholdsvis anparts- og aktieselskaber.

JMP-vink: “Analyze” -> “Fit Y by X”. Placér Dækningsgrad (i %) under “Y, Response”, Selskabsform

under “X, Factor”. Vælg “Means and Std Dev” via den røde trekant.

Gør rede for sammenhængen mellem de beregnede middelværdier og standardafvi-gelser og de optegnede box plots.



Vi ser af boxplots der er en del outliers

Middelværdien og standardafvigelse i stikprøven ligger højere for ApS end for A/S.

b). Gør rede for om der er statistisk belæg for en nulhypotese om, at den forventede dæk-ningsgrad er højst 10% større for anparts- end for aktieselskaber (brug = 5% som signifikansniveau). Angiv i den forbindelse antagelser, hypoteser, signifikanssandsyn-lighed samt teknisk og let forståelig konklusion.

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Means”. Vælg “Raw

Data”.

|  |  |
| --- | --- |
| Test 2 Middelværdier ens varians, tosidet alternativ hypotese | |
| Forudsætninger | Det forudsættes at varianserne er uens for de 2 populationer. Vi gennemfører en Unpooled t-test  Forudsætning for test for varianshomogenitet er at begge populationer er normalfordelte. |
| Hypotese |  |
| Teststørrelse | 1.6285 |
| P-værdi/signifikanssandsynlighed | 5.18% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/ signifikanssandsynligheden 5.18% er større end 5% signifikansniveauet afviser vi ikke nulhypotesen. Bemærk konklusionen er niveaufølsom |
| Ikke-teknisk konklusion | Den forventede dækningsgrad er højst 10% større for anparts- end for aktieselskaber. |
| Ved brug af calculator: | |

c). Gør rede for om konklusionen i delspg. b) ville ændre sig, såfremt datamaterialet kun bestod af 100 aktie- og 100 anpartsselskaber (og alle andre størrelser antages uæn-drede).

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Means”. Vælg “Sum-

mary Statistics”.

|  |  |
| --- | --- |
| Test 2 Middelværdier ens varians, tosidet alternativ hypotese | |
| Forudsætninger | Det forudsættes at varianserne er uens for de 2 populationer. Vi gennemfører en Unpooled t-test  Forudsætning for test for varianshomogenitet er at begge populationer er normalfordelte. |
| Hypotese |  |
| Teststørrelse | 0.558 |
| P-værdi/signifikanssandsynlighed | 28.88% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/ signifikanssandsynligheden 28.88% er større end 5% signifikansniveauet afviser vi ikke nulhypotesen. |
| Ikke-teknisk konklusion | Den forventede dækningsgrad er højst 10% større for anparts- end for aktieselskaber. |
| Ved brug af calculator: | |

Gør rede for om konklusionen i delspg. b) ville ændre sig, såfremt datamaterialet kun bestod af 50 aktie- og 50 anpartsselskaber (og alle andre størrelser antages uændre-de). JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Means”. Vælg

“Summary Statistics”.

|  |  |
| --- | --- |
| Test 2 Middelværdier ens varians, tosidet alternativ hypotese | |
| Forudsætninger | Det forudsættes at varianserne er uens for de 2 populationer. Vi gennemfører en Unpooled t-test  Forudsætning for test for varianshomogenitet er at begge populationer er normalfordelte. |
| Hypotese |  |
| Teststørrelse | 0.3946 |
| P-værdi/signifikanssandsynlighed | 34.7% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/ signifikanssandsynligheden 34.7% er større end 5% signifikansniveauet afviser vi ikke nulhypotesen. |
| Ikke-teknisk konklusion | Dækningsgraden er højst 10% højere for ApS end for A/S. |
| Ved brug af calculator: | |

Gør rede for intuitionen bag, hvordan konklusionen i det ovenstående ændres i takt med at størrelsen af datamaterialerne ændres.

Da teststørrelsen bliver mindre når stikprøvestørrelsen bliver mindre, får vi ikke nok beviser mod nulhypotesen.

3/6

HD Dataanalyse, efterår 2018 Opgaveark 7



Baggrund og databeskrivelse (opgave 3)

Denne opgave beskæftiger sig med hvilke faktorer, der er afgørende for, om ansatte er til-fredse med deres job. Opgaven tager udgangspunkt i filen Jobtilfredshed.jmp, hvis indhold er beskrevet på opgaveark 4.

Opgave 3

a). Beregn et 95%-konfidensinterval for forskellen i sandsynligheden for at være tilfreds med sit job (Jobtilfredshed) blandt offentligt og privat ansatte (Offentlig/privat) og giv en fortolkning af intervallet.

JMP-vink: “Analyze” -> “Fit Y by X”. Placér Jobtilfredshed under “Y, Response”, Oﬀentlig/privat under

“X, Factor”. Vælg “Two Sample Test for Proportions” via den røde trekant.



Vi kan med 95% sikkerhed sige at forskellen mellem andelen af tilfredse offentlige ansatte minus andelen af tilfredse private ansatte ligger mellem -0.459% og 9.24%.

Beregn et 99%-konfidensinterval for forskellen i sandsynligheden for at være utilfreds med sit job (Jobtilfredshed) blandt oﬀentligt og privat ansatte (Oﬀentlig/privat) og

giv en fortolkning af intervallet.

JMP-vink: “Analyze” -> “Fit Y by X”. Placér Jobtilfredshed under “Y, Response”, Oﬀentlig/privat under “X, Factor”. Vælg “Two Sample Test for Proportions” via den røde trekant. Vælg “Set Level” via den røde trekant.



Vi kan med 99% sikkerhed sige at forskellen mellem andelen af tilfredse offentlige ansatte minus andelen af tilfredse private ansatte ligger mellem -10.77% og 1.98%.

b). Gør rede for om der er statistisk belæg for en nulhypotese om, at jobtilfredsheden er den samme blandt privat og offentligt ansatte (brug = 5% som signifikansniveau). Angiv i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Proportions”. Vælg

“Summary Statistics”.

|  |  |
| --- | --- |
| Test af 2 andele, tosidet alternativ hypotese | |
| Forudsætninger | skal være større end 9  skal være større end 9 |
| Hypotese |  |
| Teststørrelse | -1.82 |
| P-værdi/signifikanssandsynlighed | 6.80% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/ signifikanssandsynligheden 6.80% er større end 5% signifikansniveauet forkaster vi ikke nulhypotesen. |
| Ikke-teknisk konklusion | Jobtilfredsheden er den samme blandt privat og offentligt ansatte. |
| Eller Pearson teststørrelsen der svarer til pooled test | |

Gør rede for om der er statistisk belæg for en nulhypotese om, at jobtilfredsheden er mindst 2% lavere blandt offentligt end blandt private ansatte (brug = 5% som

signifikansniveau).

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Proportions”. Vælg

“Summary Statistics”.

**mindst 2% lavere** er det samme som **højst -2%**

Tænk på det på denne måde:

Hvis jobtilfredsheden fx er 2%, 3%, 4% eller 5% lavere blandt offentligt end blandt private ansatte så er nulhypotesen sand.

Så bliver nulhypotesen fra teksten:

Dette er en supersvær opgave. Vi kan gøre det lidt nemmere, hvis vi opskriver nulhypotesen med den største andel først (så passer testet også med JMP’s value ordering), det vil betyde at forskellen bliver positiv dvs. 2% i stedet for -2%.

Vi starter med:

Vi ganger nu -1 ind i uligheden, det må vi godt når vi husker at vende ulighedstegnet:

Vi flytter rundt på leddene:

Vi kan få samme hypotese, som i den vejledende løsning igen, ved at gange med igennem med -1 (husk så vender vi ulighedstegnet).

|  |  |
| --- | --- |
| Test af 2 andele, tosidet alternativ hypotese | |
| Forudsætninger | skal være større end 9  skal være større end 9 |
| Hypotese | Eller vi kan skrive hypotesen som i den vejledende løsning vi ganger igennem med -1 i begge uligheder husk ulighedstegnet vendes så: |
| Teststørrelse | 2.64 |
| P-værdi/signifikanssandsynlighed | 0.41% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/signifikanssandsynligheden 0.41% er mindre end 5% signifikansniveauet, forkaster vi nulhypotesen. |
| Ikke-teknisk konklusion | Hvis vi formulerer ud fra spørgsmålet, kan vi skrive:  Jobtilfredsheden er ikke mindst 2% lavere blandt offentligt end blandt privat ansatte.  Eller ud fra alternativ-hypotesen:  Jobtilfredsheden er under 2% højere blandt private end blandt offentligt ansatte.  Eller kortere ud fra alternativ-hypotesen:  Jobtilfredsheden under 2% højere blandt private ansatte. |
|  | |

c). Gør rede for om der er statistisk belæg for en nulhypotese om, at jobtilfredsheden er mindst 1% større blandt mænd end blandt kvinder (Køn) (brug = 5% som signifi-kansniveau). Angiv i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Hypothesis Test for Two Proportions”. Vælg “Summary Statistics”.

|  |  |
| --- | --- |
| Test af 2 andele, tosidet alternativ hypotese | |
| Forudsætninger | skal være større end 9  skal være større end 9 |
| Hypotese |  |
| Teststørrelse | -1.709 |
| P-værdi/signifikanssandsynlighed | 4.37% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien/signifikanssandsynligheden 4.37% er mindre end 5% signifikansniveauet, forkaster vi nulhypotesen. Bemærk konklusionen er niveaufølsom |
| Ikke-teknisk konklusion | Jobtilfredsheden er ikke mindst 1% større blandt mænd. Eller jobtilfredsheden er mindre end 1% større blandt mænd. |
|  | |



