HD Dataanalyse, efterår 2018 Opgaveark 5



Baggrund og databeskrivelse (opgave 1-2)

Denne opgave ser nærmere på forskellige regnskabsnøgletal for et udsnit af små, mellem-store og store danske virksomheder, der har haft (års)regnskabsafslutning i løbet af kalen-deråret 2008.

Opgaven tager udgangspunkt i datamaterialet i filen Experian2008.jmp, der for hver virk-somhed indeholder forskellige virksomhedskarakteristika samt information om en række regnskabs- og nøgletal.

Opgave 1 (egenkapitalens forrentning)

a). Tegn et histogram over fordelingen af egenkapitalens forrentning blandt de betrag-tede virksomheder (Egenkapitalens forrentning (i %)).

JMP-vink: “Analyze” -> “Distribution”



Estimér middelværdi og standardafvigelse i den normalfordeling, der bedst beskriver virksomhedernes egenkapitalforrentning.

JMP-vink: “Analyze” -> “Distribution”

er den normalfordeling, der bedst beskriver virksomhedernes egenkapitalforrentning.

b). Beregn et 95%-konfidensinterval for middelværdien i fordelingen af virksomhedernes egenkapitalforrentning og giv en fortolkning af intervallet.

JMP-vink: “Analyze” -> “Distribution”



Vi kan med 95% sikkerhed sige at den gennemsnitlige forrentning af egenkapitalens forrentning i populationen ligger mellem -0.55% og 3.38%.

Gør rede for hvordan konfidensintervallet ændres, såfremt virksomhederne i årene fremover opnår en mere stabil egenkapitalforrentning, således at standardafvigelsen af egenkapitalens forrentning falder til det halve (og alle andre størrelser antages uændrede).

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Confidence Interval for One Mean”



Konfidensintervallet bliver smallere, såfremt virksomhederne i årene fremover opnår en mere stabil egenkapitalforrentning, således at standardafvigelsen σ af egenkapitalens forrentning falder til det halve 61. Dette skyldes der ikke vil være så meget variation i egenkapitalforrentningen.

c). Gør rede for hvordan konfidensintervallet i delspg. b) ændres, såfremt intervallet kun baseres på halvt så mange virksomheder, som i delspg. b) (og alle andre størrelser antages uændrede).

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Confidence Interval for One Mean”





Konfidensintervallet er bredere hvis antallet af observationer er mindre. Vi bruger standardfejlen (standard error of the mean SE) til at bestemme konfidensintervallet, her dividerer vi med n antallet af observationer n:

HD Dataanalyse, efterår 2018 Opgaveark 5



En række virksomheder har en forholdsvis ekstrem egenkapitalforrentning i 2008, og disse outlier observationer påvirker i meget høj grad estimationen af normalfordelingens para-metre. I den resterende del af opgave 1 begrænser vi os derfor til at se på virksomheder, hvis egenkapitalforrentning ligger mellem -100% og +100%.

JMP-vink: “Rows” -> “Data Filter”

Vi får ved data view 12543 observationer når vi fjerner outliers

d). Estimér middelværdi og standardafvigelse i den normalfordeling, der bedst beskriver egenkapitalens forrentning.

JMP-vink: “Analyze” -> “Distribution”



er den den normalfordeling, der bedst beskriver virksomhedernes egenkapitalforrentning. Vi får væsentlig højere gennemsnitlig egenkapitalforrentning, og væsentlig lavere variation når vi fjerner outliers.

Beregn medianen i den fundne normalfordeling, og giv en fortolkning af den.

JMP-vink: “Analyze” -> “Distribution”



Medianen 9.17 fortæller 50% af virksomhederne har en egenkapital forrentning på 9.17% eller derunder. Bemærk medianen har ikke ændret sig meget efter vi fjernede outliers i forhold til gennemsnittet.

e). Beregn et 90%-konfidensinterval for middelværdien i fordelingen af egenkapitalens forrentning.

JMP-vink: “Analyze” -> “Distribution”



Vi kan med 90% sikkerhed sige at den gennemsnitlige forrentning af egenkapitalens i populationen ligger mellem 11.37% og 12.51%.

Gør rede for hvordan konfidensintervallet ændres, såfremt konfidensniveauet i stedet ændres til 95%.

JMP-vink: “Analyze” -> “Distribution”



Vi kan med 95% sikkerhed sige at den gennemsnitlige forrentning af egenkapitalens i populationen ligger mellem 11.26% og 12.62%. Højere sikkerhed 95% giver ringere præcision dvs. bredere konfidensinterval.

f). Beregn hvor mange virksomheder analysen skal baseres på, såfremt den forventede egenkapitalforrentning med 95% sandsynlighed skal kunne bestemmes med en usik-kerhed på plus/minus 0,5%.

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Sample Size for Confidence Intervals”



Vi skal da have en stikprøvestørrelse på mindst 22931.

Opgave 2 (ordinært resultat)

Opgave 2 omhandler udelukkende virksomheder med en soliditetsgrad på mellem 0% og 50%.

JMP-vink: “Rows” -> “Data Filter”

a). Tegn et histogram over fordelingen af det ordinære resultat blandt de betragtede virk-somheder (Ordinært resultat (i 1.000 kr) ).

JMP-vink: “Analyze” -> “Distribution”



Estimér middelværdi og standardafvigelse i den normalfordeling, der bedst beskriver virksomhedernes ordinære resultat.

JMP-vink: “Analyze” -> “Distribution”



er den den normalfordeling, der bedst beskriver virksomhedernes ordinære resultat.

2/5

HD Dataanalyse, efterår 2018 Opgaveark 5



b). Beregn et 80%-konfidensinterval for middelværdien i fordelingen af virksomhedernes ordinære resultat.

JMP-vink: “Analyze” -> “Distribution”

**Confidence Intervals**

| **Parameter** | **Estimate** | **Lower CI** | **Upper CI** | **1-Alpha** |
| --- | --- | --- | --- | --- |
| Mean | 877.8787 | 277.721 | 1478.036 | 0.800 |
| Std Dev | 43569.23 | 43149.93 | 43998.77 | 0.800 |

Vi kan med 80% sikkerhed sige at virksomhedernes gennemsnitlige ordinære resultat.

i populationen ligger mellem 277.721 DKK og 1.478.036 DKK.

c). Gør rede for antagelserne bag beregningen af intervallet i delspg. b), og diskutér i hvilket omgang antagelserne virker rimelige i forhold til det givne datamateriale.

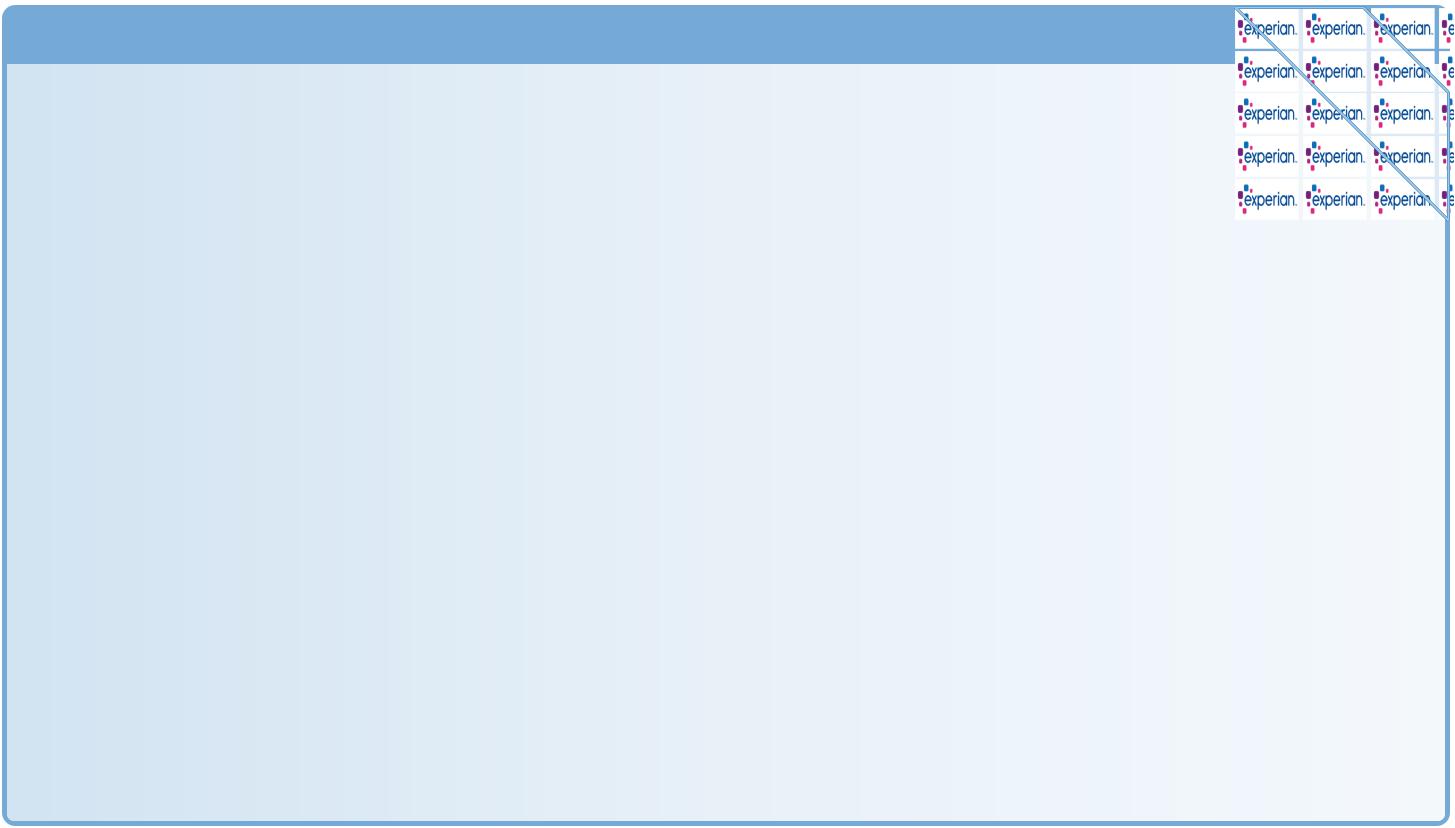
Stikprøven skal være repræsentativ i forhold til populationen, eftersom vi har filtreret hårdt vha. soliditetsgrad kan dette være tvivlsomt.

d). Gør rede for hvordan konfidensintervallet i delspg. b) ændres, såfremt alle de betrag-tede virksomheder havde opnået et ordinært resultat, der var 200.000 kr. lavere end det, der fremgår af datamaterialet.

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Confidence Interval for One Mean”



80%-konfidensinterval (hvis overskud 200.000 kr. lavere): [78; 1*:*278]



Datakilde (opgave 1-2)

Datamaterialet i ovenstående opgaver er venligst stillet til rådighed af data- og analysevirksomheden ExperianTM [(experian.dk)](http://www.experian.dk/).

Virksomheden giver følgende beskrivelse af sig selv:

Vi er verdens førende leverandør af data, analyser og statistiske modeller. Vi hjælper virksom-heder med at styre kreditrisiko, forebygge svindel, målrette markedsføring og automatisere beslutningsprocesser. Vi hjælper også privatpersoner med at kende deres kreditværdighed og beskytte sig mod identitetstyveri. I 2015 udnævnte Forbes Magazine os til at være én af verdens mest innovative virksomheder.

Vi beskæftiger omkring 17.000 mennesker i 37 lande. Vores hovedkontor ligger i Dublin, Irland, med operative hovedkvarterer i Nottingham (UK), Californien (USA) og São Paulo (Brasilien).

DISCLAIMER:

Datamaterialet i nærværende opgaver må udelukkende anvendes i forbindelse med undervisningen i faget Dataanalyse på CBS. Datamaterialet er fortroligt og må ikke

videredistribueres eller oﬀentliggøres andet end overfor studerende, der følger faget Dataanalyse på CBS, hverken skriftligt, mundtligt eller på anden vis, og ej heller de

resultater og indsigter, der måtte komme ud af arbejdet med datamaterialet.





HD Dataanalyse, efterår 2018 Opgaveark 5



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). Estimér sandsynligheden for at være tilfreds med sit job i en passende valgt binomial-fordeling, der beskriver fordelingen af de adspurgtes jobtilfredshed (Jobtilfredshed).

JMP-vink: “Analyze” -> “Distribution”



Vi ser estimaet for p dvs.

Beregn et 95%-konfidensinterval for sandsynligheden for at være tilfreds og giv en fortolkning af intervallet.

JMP-vink: “Analyze” -> “Distribution”



Vi kan med 95% sikkerhed sige at andelen af tilfredse i populationen ligger mellem 88.47% og 93.34%

b). Beregn et 95%-konfidensinterval for sandsynligheden for at være utilfreds med sit job og giv en fortolkning af intervallet.

JMP-vink: “Analyze” -> “Distribution”



Vi kan med 95% sikkerhed sige at andelen af utilfredse i populationen ligger mellem 6.66% og 11.53%

c). Gør rede for hvordan konfidensintervallet i delspg. a) ændres, såfremt konfidensni-veauet i stedet ændres til 99%.

JMP-vink: “Analyze” -> “Distribution”

Konfidensintervallet bliver bredere da vi får øget sikkerhed nu 99%

Vi kan med 99% sikkerhed sige at andelen af utilfredse i populationen ligger mellem 87.48% og 93.90%



Gør rede for hvordan konfidensintervallet i delspg. a) ændres, såfremt antallet af ads-purgte i stedet ændres til 100 personer (og alle andre størrelser antages uændrede).

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Confidence Interval for One Proportion”



Konfidensintervallet bliver bredere da stikprøven bliver mindre.

Vi kan med 95% sikkerhed sige at andelen af tilfredse i populationen ligger mellem 85.4% og 96.6%

d). Beregn hvor mange personer, der skal udspørges om deres jobtilfredshed, for at sand-synligheden for at være tilfreds med sit job kan bestemmes med en usikkerhed på plus/minus 5%.

JMP-vink: “Help” -> “Sample Data” -> “Calculators” -> “Sample Size for Confidence Intervals”



For at få en ønsket fejlmargin på 5 procentpoint skal stikprøvestørrelsen være på 126



For at få en ønsket fejlmargin på 1 procentpoint skal stikprøvestørrelsen være på 3147

Beregn hvor mange personer, der skal udspørges om deres jobtilfredshed, for at sand-synligheden for at være tilfreds med sit job med 95% sandsynlighed kan bestemmes med en usikkerhed på plus/minus 1%.



4/5

HD Dataanalyse, efterår 2018 Opgaveark 5



Kortfattede talløsninger

Opgave 1

a): ˆ = 1*;*41%; ˆ = 122*;*23%

b): 95%-konfidensinterval: [ 0*;*55%; 3*;*38%]

95%-konfidensinterval hvis standardafvigelsen halveres: [0*;*43%; 2*;*39%]

NB: Resultatet kan afvige på decimalerne afhængig af præcis hvilke tal, der indtastes i JMP

c): 95%-konfidensinterval hvis antal observationer halveres: [ 1*;*37%; 4*;*19%]

NB: Resultatet kan afvige på decimalerne afhængig af præcis hvilke tal, der indtastes i JMP

|  |  |
| --- | --- |
| d): ˆ = 11*;*94%; ˆ = | 38*;*63% |
| Median = 11*;*94% |  |

e): 90%-konfidensinterval: [11*;*37%; 12*;*51%]

95%-konfidensinterval: [11*;*26%; 12*;*62%]

f): Antal observationer: 22.931

NB: Resultatet kan afvige på decimalerne afhængig af præcis hvilke tal, der indtastes i JMP

Opgave 2

a): ˆ = 877*;*8; ˆ = 43*:*569*;*23

b): 80%-konfidensinterval: [277*;*72; 1*:*478*;*04]

d): 80%-konfidensinterval (hvis overskud 200.000 kr. lavere): [77*;*72; 1*:*278*;*04]

Opgave 3

a): *p*ˆ = 91*;*2%

95%-konfidensinterval: [88*;*5%; 93*;*3%]

b): 95%-konfidensinterval (for utilfreds): [6*;*7%; 11*;*5%]

c): 99%-konfidensinterval (for tilfreds): [87*;*5%; 93*;*9%]

95%-konfidensinterval (ved 100 observationer): [85*;*4%; 96*;*6%]

NB: Resultatet kan afvige på decimalerne afhængig af præcis hvilke tal, der indtastes i JMP

d): Antal observationer: 3.147

NB: Resultatet kan afvige på decimalerne afhængig af præcis hvilke tal, der indtastes i JMP



5/5