

Copenhagen Business School
Dataanalyse, HD 1. del

ÆKSEMPEL PÅ EKSAMENSOPGAVE

4 timers skriftlig eksamen

Opgaven består af i alt **31** delspørgsmål fordelt på **10** opgaver. Ved bedømmelsen vægtes alle opgavens delspørgsmål ligeligt med undtagelse af delspørgsmål 10 a), der vægtes dobbelt. Vægtene skal alene opfattes som vejledende med henblik på disponering af tiden ved besvarelsen af eksamensopgaven.

NB:

- Hvis værktøjslinjen i JMP ikke fremgår af programmet, kan den genskabes v.hj.a. <Ctrl>-<K> -> “Windows Specific” -> “Auto-hide menu and toolbars” -> “never”
- Husk at medtage relevant JMP-output i opgavebesvarelsen (vedlæg ikke alt output men kun de dele, der refereres til i opgavebesvarelsen)
- Husk at JMP-output kun tæller med i bedømmelsen, såfremt det er eksplicit kommenteret/refereret.



VIRKSOMHEDSNØGLETAL

BAGGRUND OG DATABESKRIVELSE

Næsten 93% af alle danske virksomheder er små virksomheder med 0-10 ansatte.¹ For sådanne virksomheder udgør årsregnskabet den vigtigste kilde til information om virksomhedens økonomiske og finansielle tilstand. For at kunne vurdere eksempelvis kreditværdighed og profitabilitet (dvs. evnen til at tjene penge) for sådanne virksomheder, er det derfor essentielt at kunne analysere indholdet af deres regnskaber.

Denne opgave ser nærmere på forskellig regnskabsinformation for et udsnit af små, mellemstore og store danske virksomheder over perioden 2007-2011 ved at analysere datamaterialet i JMP-filen *Experian.jmp*.

For hver virksomhed og hvert regnskabsår indeholder datamaterialet en række virksomhedskarakteristika, informationer om en række regnskabs- og nøgletal samt information om, hvorvidt virksomheden gik konkurs eller ej i det betragtede regnskabsår.

Eksamensopgaven skal besvares på baggrund af datamaterialet i *Experian.jmp*, hvis indhold er nærmere beskrevet i tabellen på de næste sider.

¹Baseret på data for år 2016 fra Danmarks Statistik, Statistikbanken, tabel FGF3.



VARIABELNAVN	FORKLARING
<i>Virksomhedens stamdata</i>	
<i>VirksomhedsID</i>	Numerisk virksomheds-id
<i>Primær branche</i>	Hvilken branche virksomheden primært er beskæftiget indenfor
<i>Postnummer</i>	Postnummer for virksomhedens hovedsæde
<i>Postdistrikt</i>	Postdistrikt for virksomhedens hovedsæde
<i>Region</i>	Hvilken region virksomhedens hovedsæde er beliggende i (med de mulige værdier "Sjælland", "Hovedstaden", "Nordjylland", "Midtjylland", "Syddanmark", "Grønland")
<i>Etableringsår</i>	Hvilket årstal virksomheden er etableret i
<i>Selskabsform</i>	Information om virksomheden er et anparts- eller aktieselskab (med de mulige værdier "APS", "A/S")
<i>Virksomhedens regnskabstal</i>	
<i>Regnskabsår</i>	Det regnskabsår regnskabs- og nøgletal stammer fra
<i>Samlet balance</i>	Værdien af virksomhedens samlede balance (i 1.000 kr.), også kaldet "virksomhedens samlede kapital".
<i>Nettoomsætning/balance</i>	Virksomhedens nettoomsætning divideret med virksomhedens samlede balance
<i>Anlægsaktiver/balance</i>	Virksomhedens anlægsaktiver divideret med virksomhedens samlede balance. Anlægsaktiver er aktiver, der er til vedvarende brug eller eje for virksomheden og er derfor aktiver, som kan være svære at afhænde hurtigt, f.eks. med det formål at skaffe kapital.
<i>Egenkapital/balance</i>	Virksomhedens egenkapital divideret med virksomhedens samlede balance. Egenkapitalen udtrykker værdien af virksomhedsejernes andel af virksomhedens samlede værdi.
<i>Hensættelser/balance</i>	Virksomhedens hensættelser divideret med virksomhedens samlede balance. Hensættelser udtrykker kapital allokeret til forventede fremtidige forpligtelser som skat, tab på debitorer m.m.
<i>Langfr:gæld/balance</i>	Virksomhedens langfristede gæld divideret med virksomhedens samlede balance. Langfristet gæld er gæld der forfalder om <i>me-</i> re end ét år.
<i>Kortfr:gæld/balance</i>	Virksomhedens kortfristede gæld divideret med virksomhedens samlede balance. Kortfristet gæld er gæld der forfalder om <i>min-</i> dre end ét år.
<i>Antal ansatte</i>	Antal ansatte i virksomheden



VARIABELNAVN (fortsat)	FORKLARING (fortsat)
<i>Virksomhedens nøgletal (alle angivet i procent)</i>	
<i>Afkastningsgrad</i>	Nøgletal der udtrykker, hvor stor en del af afkastet af virksomhedens samlede kapital, der skyldes virksomhedens primære aktiviteter. En højere værdi udtrykker - alt andet lige - en mere profitabel virksomhed.
<i>Likviditetsgrad</i>	Nøgletal der udtrykker virksomhedens evne til at honorere sine kortfristede gældsforpligtelser. En højere værdi udtrykker - alt andet lige - en mere finansielt velpolstret virksomhed.
<i>Soliditetsgrad</i>	Nøgletal der udtrykker, hvor stor en del af virksomhedens kapital, der udgøres af virksomhedsejernes egen finansiering (modsat lånefinansiering).
<i>Egenkapitalens forrentning</i>	Nøgletal der udtrykker, hvor meget virksomhedsejernes kapital forrentes med.
<i>Konkursstatus</i>	
<i>Status</i>	Information om virksomheden er aktiv eller er gået konkurs (med de mulige værdier "Aktiv", "Konkurs")
<i>Konkursbehandlingsdato</i>	Dato for hvornår virksomheden er taget under konkursbehandling, såfremt den er gået konkurs (dvs. såfremt <i>Status</i> = Konkurs)



OPGAVE 1 (ETABLERINGSÅR)

For at få overblik hvor længe de betragtede virksomheder har eksisteret, ser vi på, i hvilket år virksomhederne er etableret, som angivet i variabelen *Etableringsår*.

- Tegn** et histogram og et boxplot over fordelingen af variabelen *Etableringsår*.
Giv en kort beskrivelse af fordelingen af variabelen *Etableringsår* på baggrund af histogrammet, boxplottet og relevante nøgletal for centrum og variation.
- Beregn** interkvartilbredden i fordelingen af variabelen *Etableringsår*.
Gør rede for om standardafvigelse eller interkvartilbredde (eller begge) er retvisende mål for variationen i fordelingen af variabelen *Etableringsår*.
- Angiv og fortolk** 5%-fraktilen i fordelingen af variabelen *Etableringsår*.
Gør rede for om flere eller færre end 5% af de betragtede virksomheder er etableret før år 1900.
- Tegn** den kumulerede fordelingsfunktion (CDF) for variabelen *Etableringsår*.
Gør rede for hvorledes den kumulerede fordelingsfunktion indikerer, at antallet af nyetablerede virksomheder har været nogenlunde konstant siden år 2000.
Vink: Begræns 1. akse i CDF-plottet til at starte i værdien 2000.
Gør rede for hvad CDF-plottet indikerer om alderen på de 20% yngste virksomheder.

OPGAVE 2 (SELKABSFORM)

Virksomhederne i datamaterialet er enten anparts- eller aktieselskaber. For at undersøge om selskabsformen har en sammenhæng med virksomhedens størrelse, ser vi på selskabsform, som angivet i variabelen *Selskabsform*, og virksomhedens størrelse målt ved dens samlede balance, som angivet i variabelen *Samlet balance*.

- Gør rede for** om der er statistisk belæg for en nulhypotese om, at den forventede andel af aktieselskaber er lig 50%. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.
- Beregn** middelværdi og median for variabelen *Samlet balance*.
Gør rede for hvad de to nøgletal indikerer om fordelingen af variabelen *Samlet balance*.
- Beregn** for hver af de to selskabsformer middelværdi og standardafvigelse for variabelen *Samlet balance*.
Gør rede for hvad de beregnede nøgletal indikerer om eventuelle ligheder/forskelle på størrelsen af aktie- og anpartsselskaber.



OPGAVE 3 (VIRKSOMHEDENS STØRRELSE)

Størrelsen af en virksomhed forventes normalt i et vist omfang at influere på virksomhedens evne til at skabe overskud. Når man vil se på virksomheders profitabilitet, er det derfor relevant at tage højde for, at de varierer i størrelse. Hvordan man opgør størrelsen af en virksomhed, er imidlertid ikke helt klart. Vi ser derfor her på forskellige mål for en virksomheds størrelse for at undersøge, om det betyder noget, hvordan man opgør virksomhedens størrelse.

- a). **Tegn** et scatterplot af variablene *Etableringsår*, *Samlet balance* og *Antal ansatte* mod hinanden.
Beregn variablenes indbyrdes korrelation og **gør rede for** om variablene ser ud til, at indeholde stort set samme information om virksomhedens størrelse.
- b). **Gør rede for** i hvilket omfang korrelationerne i delspg. a) kan siges at udtrykke en kausal sammenhæng.

OPGAVE 4 (AFKASTNINGSGRAD)

Ved hjælp af regnskabstallene er det muligt bl.a. at undersøge virksomhedernes profitabilitet, der f.eks. kan måles ved nøgletallet afkastningsgrad.

- a). **Estimér** middelværdi og standardafvigelse i den normalfordeling, der bedst beskriver fordelingen af variablen *Afkastningsgrad*.
- b). **Tegn** et histogram over fordelingen af variablen *Afkastningsgrad*.
Gør rede for (på baggrund af histogrammet) om normalfordelingen ser ud til at give en rimelig beskrivelse af fordelingen af variablen *Afkastningsgrad*.
- c). **Beregn** på baggrund af den i delspg. a) beregnede normalfordeling, sandsynligheden for at en virksomhed har en positiv afkastningsgrad.

OPGAVE 5 (AFKASTNINGSGRAD OG REGNSKABSÅR)

Virksomheders profitabilitet varierer ofte fra år til år, bl.a. som følge af ændringer i de generelle økonomiske konjunkturer. Det er derfor interessant at undersøge, om der ser ud til at være systematiske forskelle i virksomhedernes profitabilitet over den betragtede periode 2007-2011.

- a). **Tegn** et boxplot over fordelingen af variablen *Afkastningsgrad* for hver værdi af variablen *Regnskabsår*.
Gør rede for hvad de tegnede boxplots indikerer om eventuelle forskelle i afkastningsgraden over tid.
- b). **Gør rede for** om der er statistisk belæg for en nulhypotese om, at den forventede afkastningsgrad er uændret over tid. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

**OPGAVE 6 (AFKASTNINGSGRAD, SAMLET ANALYSE)**

En oplagt anvendelse af regnskabstallene er til at undersøge hvilke faktorer, der er afgørende for virksomhedernes profitabilitet. Vi ser derfor på en lineær regressionsmodel med variabelen *Afkastningsgrad* som responsvariabel og variablene *Region*, *Etableringsår*, *Selskabsform*, *Regnskabsår*, *Nettoomsætning/balance*, *Anlægsaktiver/balance*, *Egenkapital/balance*, *Hensættelser/balance*, *Langfr.gæld/balance*, *Kortfr.gæld/balance*, *Likviditetsgrad*, *Soliditetsgrad*, *Status* som forklarende variable.

- Reducér** den angivne regressionsmodel til kun at indeholde signifikante variable.
Vink: Fjern insignifikante variable fra modellen én ad gangen ved hele tiden at fjerne den mest insignifikante variabel.
- Angiv** estimater for regressionskoefficienter hørende til variablene i den reducerede model og **giv** en fortolkning af estimaterne.
- Gør rede for** om R^2 -værdien er størst for den oprindelige eller den reducerede model, og **forklar** hvorfor.
- Beregn** et 95%-konfidensinterval for betydningen af selskabsform og **giv** en fortolkning af intervallet.
- Gør rede for** om det betyder mest for den forventede afkastningsgrad, om variabelen *Nettoomsætning/balance* stiger med 0,01 eller om variabelen *Kortfr.gæld/balance* falder med 0,01.
- Beregn** forskellen i den forventede afkastningsgrad, såfremt virksomheden er beliggende i region Hovedstaden frem for i region Syddanmark.
Beregn forskellen i den forventede afkastningsgrad, såfremt virksomheden er beliggende i region Hovedstaden frem for på Grønland.

OPGAVE 7 (STATUS OG SELSKABSFORM)

Et vigtigt element af informationen i datafilen er information om, hvorvidt en given virksomhed er gået konkurs i det givne regnskabsår eller ej, som angivet i variabelen *Status*. I den forbindelse er det blandt andet relevant at se på, om der er forskel i andelen af konkursramte virksomheder blandt anparts- og aktieselskaber.

- Opstil** en krydstabel for de to variable *Selskabsform* og *Status*.
Beregn sandsynligheden for at en tilfældigt valgt konkursramt virksomhed er et anpartsselskab.
Beregn sandsynligheden for at en tilfældigt valgt virksomhed er konkursramt og er et anpartsselskab.
- Estimer** for hver af de to selskabsformer (anparts- hhv. aktieselskab) andelen af konkursramte virksomheder.



Gør rede for om der er statistisk belæg for en nulhypotese om, at der ikke er forskel på andelen af konkursramte virksomheder blandt de to selskabsformer. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

- c). **Beregn** et 95%-konfidensinterval for forskellen i den forventede andel konkursramte virksomheder blandt de to selskabsformer.

Gør rede for hvorledes konfidensintervallet – alt andet lige – ændres, såfremt antallet af virksomheder i datamaterialet øges.

OPGAVE 8 (STATUS OG REGION)

En virksomheds geografiske placering kan i mange tilfælde være afgørende for dens langsigtede overlevelsesmuligheder. Det er derfor relevant at se på, om der er en sammenhæng mellem virksomhedernes geografiske placering, og i hvilket omfang de går konkurs.

- a). **Opstil** en krydstabel for de to variable *Region* og *Status*.

Gør rede for hvilken af de seks regioner, der har den største andel af konkursramte virksomheder.

- b). **Beregn** sandsynligheden for at en tilfældigt valgt virksomhed er konkursramt, givet at vi kun ser på virksomheder i region Nordjylland og region Midtjylland.

- c). **Gør rede for** om der er statistisk belæg for en nulhypotese om, at der er uafhængighed mellem hvilken region, virksomheden er beliggende i, og sandsynligheden for at være konkursramt. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

Gør rede for hvad der ud fra krydstabellen ser ud til at være forklaringen på, at nulhypotesen accepteres/forkastes.

- d). **Gør rede for** om konklusionen i delspg. c) ændres, såfremt vi ser bort fra virksomheder i region Hovedstaden. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

OPGAVE 9 (STATUS OG KORTFRISTET GÆLD)

Udover ydre faktorer som selskabsform og regional placering er det interessant at se på, i hvilket omfang virksomhedernes finansielle struktur påvirker i hvilket omfang, de går konkurs. Vi ser derfor på, om omfanget af optagelse af kortfristet gæld, som angivet i variablen *Kortfr.gæld/balance*, influerer på risikoen for at gå konkurs.

- a). **Tegn** et boxplot over fordelingen af variablen *Kortfr.gæld/balance* for henholdsvis aktive og konkursramte virksomheder.



Gør rede for hvad de to boxplots indikerer om eventuelle forskelle i omfanget af kortfristet gæld blandt de to typer virksomheder (aktive hhv. konkursramte).

- b). **Beregn** for hver af de to typer virksomheder middelværdi og standardafvigelse af variabelen *Kortfr.gæld/balance*.

Gør rede for hvorledes de beregnede nøgletal indikerer de samme tendenser som de i delspg. a) optegnede box plots.

- c). **Gør rede for** om der er statistisk belæg for en nulhypotese om, at den forventede kortfristede gæld er ens for de to typer virksomheder. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

OPGAVE 10 (STATUS OG AFKASTNINGSGRAD)

- a). **Gennemfør** en passende valgt analyse som undersøger, hvorvidt datamaterialet understøtter en påstand om, at der er forskel i profitabiliteten, som målt ved virksomhedens afkastningsgrad, på aktive og konkursramte virksomheder.

DATAKILDE

Datamaterialet i ovenstående opgaver er venligst stillet til rådighed af data- og analysevirksomheden **ExperianTM** (experian.dk). Virksomheden giver følgende beskrivelse af sig selv:

Vi er verdens førende leverandør af data, analyser og statistiske modeller. Vi hjælper virksomheder 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 opgave må udelukkende anvendes i forbindelse med 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.



VEJLEDENDE LØSNING

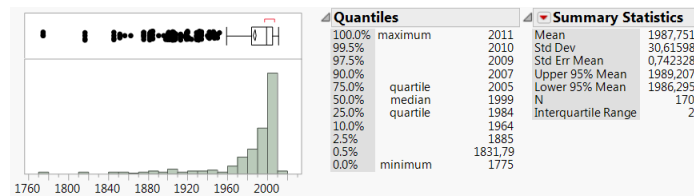
Bemærk at nedenstående løsning...

- udelukkende er **vejledende**, og at der i visse tilfælde kan være mere end én måde at besvare en given opgave på.
- visse steder har anført teoretiske formler sammen med det aflæste JMP-output for at fremme læsbarheden. Det er ikke et krav, at man gør noget tilsvarende i sin eksamensbesvarelse. Det er tilstrækkeligt at anføre de relevante tal direkte fra JMPs output samt inkludere det relevante output i besvarelsen.
- visse steder kommenterer på anførte antagelser. Det er ikke et krav, at man kommenterer på antagelser, medmindre det eksplicit efterspørges i opgaveteksten.
- til hver opgave i feltet ► *Løsning i JMP* indeholder et klikbart link til en kort video, der viser de nødvendige beregninger i JMP.



OPGAVE 1

1A): Histogrammet over fordelingen af etableringsår viser, at der er tale om en meget venstreskæv fordeling, hvor enkelte virksomheder er meget gamle med etablering helt tilbage til år 1775. Sådanne virksomheder må umiddelbart betragtes som outliers.



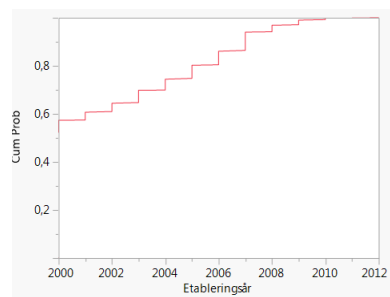
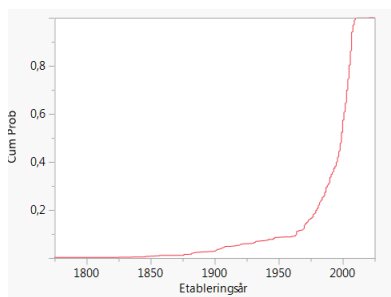
Venstreskævheden af fordelingen fremgår også delvist af boxplottet, som blandt andet viser, at middelværdien på (år) 1988 er mindre end medianen på (år) 1999, hvilket alt andet lige er en indikation på venstreskævhed. At der er tale om en del relativt gamle virksomheder, ses eksempelvis af, at 2,5% af virksomhederne er etableret i år 1885 eller før. Ligeledes ses det af, at standardafvigelsen, som er følsom overfor outliers, er påvirket af andelen af gamle virksomheder, idet standardafvigelsen er større (30,6) end interkvartilbredden (på 2005-1984 = 21), der ikke er følsom overfor outliers.

1B): Interkvartilbredden er (jf. ovenfor) 21, svarende til at de 50% "midterste" virksomheder (målt på deres alder) varierer med op til 21 år i alder. På grund af den konstaterede venstreskævhed af fordelingen af etableringsår (jf. også bemærkningerne ovenfor) er interkvartilbredden et fint og retvisende mål for variationen i fordelingen, mens standardafvigelsen i nogen grad er påvirket af fordelingsens venstreskævhed (idet den er næsten 50% større end interkvartilbredden). Interkvartilbredden er derfor umiddelbart det foretrukne mål for variationen i fordelingen.

1C): 5%-fraktilen er givet som (år) 1913, svarende til at 5% af virksomhederne i datamaterialet er etableret i år 1913 eller før. Eftersom der er virksomheder, som er etableret mellem år 1900 og år 1913 (hvilket f.eks. ses af histogrammet ovenfor), er der dermed færre end 5% af virksomhederne, som er etableret før år 1900.

Quantiles				
Quantile	Estimate	Lower 95%	Upper 95%	Actual Coverage
5%	1913	1905	1929	95,44
Smoothed Empirical Likelihood Quantiles				
Quantile	Estimate	Lower 95%	Upper 95%	
5%	1913,02	1904,89	1928,57	

1D): Den kumulerede fordelingsfunktion for virksomhedernes etableringsår viser, at antallet af nye virksomheder har været nogenlunde konstant siden år 2000, idet hældningen på CDF-kurven er nogenlunde konstant fra år 2000 og fremefter (det ses nemmest ved at "zoome ind" og kun se på perioden fra år 2000 og fremefter). Endvidere viser fordelingsfunktionen, at de ca. 80% ældste virksomheder er etableret i år 2005 eller tidligere, og dermed at de ca. 20% yngste virksomheder er etableret i år 2006 eller senere.





OPGAVE 2

2A): Test af nulhypotese om at den forventede andel af aktieselskaber er lig 50%:

Antagelser: Observationerne af de enkelte virksomheders selskabsform er indbyrdes uafhængige. Endvidere er det antallet af observationer af hver af de to mulige selskabsformer større end 15. Sidstnævnte er opfyldt, da der er henholdsvis 847 og 854 observationer af de to selskabsformer.

Hypoteser: $H_0 : p = 0,50$

$H_a : p \neq 0,50$

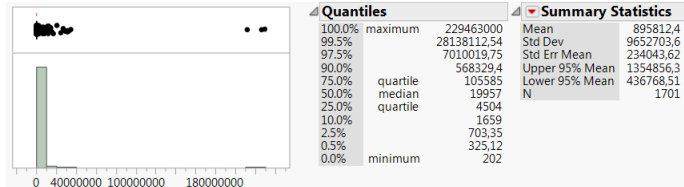
Signifikanssssh.: $p = 0,8652$

Tekn. konklusion: Signifikanssandsynligheden er større end 0,05 og dermed kan H_0 ikke forkastes.

Let konklusion: Vi kan ikke afvise hypotesen om, at andelen af aktieselskaber er lig 50%. M.a.o.: Vi kan ikke afvise hypotesen om, at der er lige mange aktie- og anpartsselskaber.

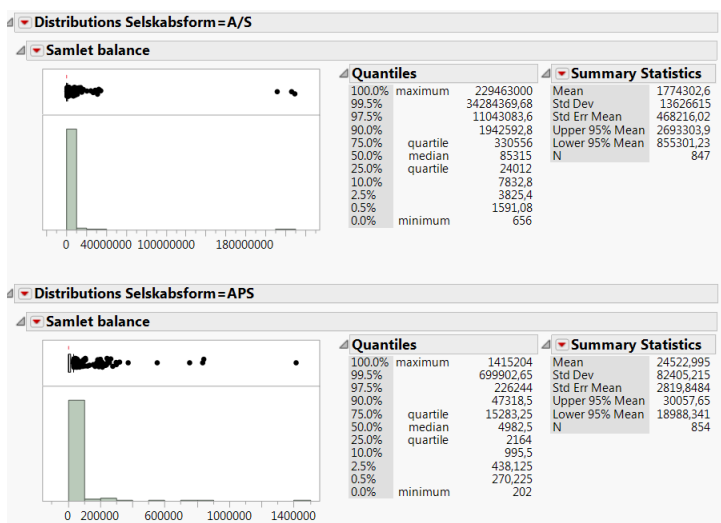
Frequencies		
Level	Count	Prob
A/S	847	0,49794
APS	854	0,50206
Total	1701	1,00000
N Missing		0
2 Levels		
Test Probabilities		
Level	Estim Prob	Hypoth Prob
A/S	0,49794	0,50000
APS	0,50206	0,50000
Test	ChiSquare	DF
Likelihood Ratio	0,0288	1
Pearson	0,0288	1
		0,8652

2B): Middelværdi og median for fordelingen af virksomhedernes samlede balance viser, at middelværdien (895,812 mio. kr.) er væsentligt større end medianen (19,957 mio. kr.), hvilket alt andet lige er en indikation på en højreskæv fordeling, svarende til at der er enkelte virksomheder, som har en markant større samlet balance end alle øvrige virksomheder i datasættet, hvilket i øvrigt også understøttes visuelt af fordelings histogram.



2c): Middelværdierne af virksomhedernes samlede balance for hver af de to selskabsformer viser, at aktieselskaber med en gennemsnitlig balance på 1.774 mio. kr. generelt er væsentlig større end anpartsselskaber, der har en gennemsnitlig balance på 25 mio. kr.

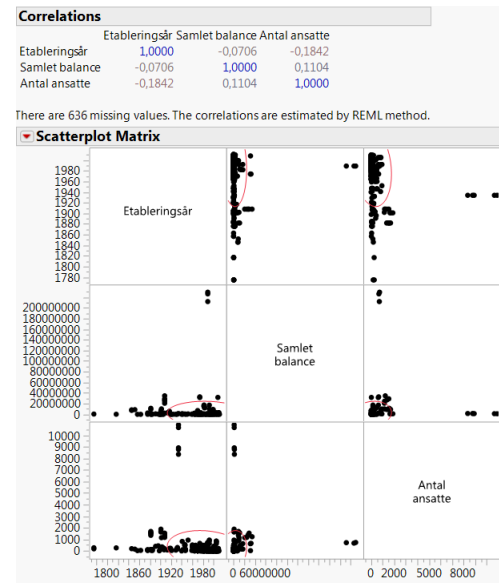
Ligeledes viser de to standardafvigelse, at variationen i de betragtede selskabers størrelse er langt større blandt aktieselskaber med en standardafvigelse på 13.627 mio. kr., mens standardafvigelsen blandt anpartsselskaber blot er på 82 mio. kr.





OPGAVE 3

3A): Scatterplottene af virksomhedernes etableringsår mod henholdsvis deres samlede balance og deres antal ansatte kan umiddelbart være en anelse svære at fortolke, fordi der er enkelte outliers som gør, at størstedelen af observationerne bliver klumpet sammen i små områder af figurerne. Korrelationen mellem etableringsår og samlet balance er $-0,07$, hvilket indikerer en meget svag negativ sammenhæng, således at nyligt etablerede virksomheder alt andet lige også er mindre (f.eks. fordi de endnu ikke har haft tid til at vokse sig store). Korrelationen mellem etableringsår og antal ansatte er $-0,18$, hvilket ligeledes indikerer en svag negativ sammenhæng, således at nyligt etablerede virksomheder alt andet lige har færre ansatte (f.eks. fordi de endnu ikke har haft tid til at vokse sig store). Korrelationen mellem samlet balance og antal ansatte er $0,11$, hvilket indikerer en svag positiv sammenhæng, således at finansielt set store virksomheder alt andet lige også har flere ansatte.



Ingen af de tre beregnede korrelationer er særligt store (i.e. de er alle forholdsvis tæt på 0), hvilket indikerer, at de tre variable etableringsår, samlet balance og antal ansatte ikke i væsentlig grad indeholder den samme information om en virksomheds størrelse.

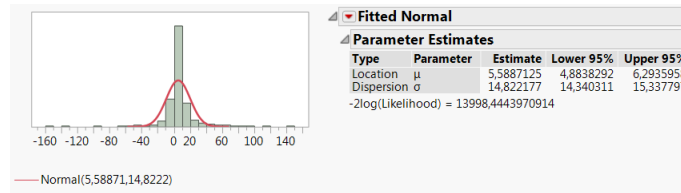
3B): Korrelation mellem to variable udtrykker ikke nødvendigvis en årsagssammenhæng (kausal sammenhæng) mellem de to variable. I tilfældet med de tre variable etableringsår, samlet balance og antal ansatte kan man dog for alle tre korrelationer argumentere for, at korrelationen i et vist omfang også udtrykker en kausal sammenhæng idet...

- en nystartet virksomhed, fordi den er nystartet, alt andet lige må forventes at have en begrænset økonomisk størrelse (korr. mellem etableringsår og samlet balance)
- en nystartet virksomhed, fordi den er nystartet, alt andet lige må forventes at have et begrænset antal ansatte (korr. mellem etableringsår og antal ansatte)
- en virksomhed med mange ansatte, fordi den har mange ansatte, alt andet lige også må have en vis økonomisk størrelse (korr. mellem antal ansatte og samlet balance)



OPGAVE 4

4A): De betragtede selskaber har en estimeret forventet afkastningsgrad på $\hat{\mu} = 5,6 \%$ med en estimeret std.afvigelse på $\hat{\sigma} = 14,8 \%$.



4B): Histogrammet over fordelingen af virksomhedernes afkastningsgrad viser, at fordelingen er nogenlunde klokkeformet, hvilket er i god overensstemmelse med normalfordelingsantagelsen. Til gengæld viser histogrammet også, at fordelingen har meget tunge haler (dvs. en del virksomheder med hhv. meget lille og meget stor afkastningsgrad), hvilket ikke stemmer så godt overens med normalfordelingsantagelsen. Der er således både noget i histogrammet, der taler for, og noget der taler imod at normalfordelingen giver en rimelig beskrivelse af fordelingen af virksomhedernes afkastningsgrad.

4C): Sandsynligheden for at en vilkårligt valgt virksomhed har en positiv afkastningsgrad er

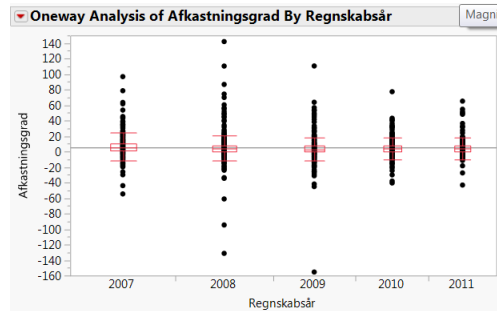
$$P(\text{Afkastningsgrad} > 0) = 1 - P(\text{Afkastningsgrad} \leq 0) = 1 - 35,3\% = 64,7\%$$

hvor sandsynligheden er beregnet i en normalfordeling med forventet værdi og std.afvigelse som angivet i delspg. a).



OPGAVE 5

5A): Boxplot over fordelingen af virksomhedernes afkastningsgrad for hvert af de betragtede regnskabsår:



Figuren viser at niveauet af afkastningsgrad, som angivet ved medianen, er nogenlunde det samme for alle regnskabsår, og at det samme gør sig gældende for variationen i afkastningsgrad som angivet ved interkvartilbredden, omend der er lidt større variation i 2008.

5B): Test af nulhypotese om at den forventede afkastningsgrad er den samme i de forskellige regnskabsår:

Oneway Anova						
Summary of Fit						
Rsquare			0,003081			
Adj Rsquare			0,00073			
Root Mean Square Error			14,81677			
Mean of Response			5,588713			
Observations (or Sum Wgts)			1701			
Analysis of Variance						
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F	
Regnskabsår	4	1150,66	287,664	1,3103	0,2639	
Error	1696	372334,11	219,537			
C. Total	1700	373484,77				
Means for Oneway Anova						
Level	Number	Mean	Std Error	Lower 95%	Upper 95%	
2007	420	6,90755	0,72298	5,4895	8,3256	
2008	389	5,34391	0,75124	3,8705	6,8174	
2009	330	4,54667	0,81564	2,9469	6,1464	
2010	292	5,22712	0,86709	3,5265	6,9278	
2011	270	5,55456	0,90172	3,7860	7,3232	

Std Error uses a pooled estimate of error variance

Antagelser: Observationerne af de enkelte virksomheders afkastningsgrader er indbyrdes uafhængige og approksimativt normalfordelte; normalfordelte indenfor hvert regnskabsår og med samme std.afvigelse for alle regnskabsår. Hvorvidt antagelsen om normalfordeling indenfor hvert regnskabsår er opfyldt kan ikke vurderes ud fra de optegnede box plots, men antagelsen om samme std.afvigelse ser umiddelbart ud til at være opfyldt, jf. boxplottene ovenfor.

Hypoteser: $H_0 : \mu_{2007} = \mu_{2008} = \mu_{2009} = \mu_{2010} = \mu_{2011}$

$H_a : \text{Mindst to af de forventede værdier } \mu_{2007}, \mu_{2008}, \mu_{2009}, \mu_{2010}, \mu_{2011} \text{ er ikke ens}$

Signifikanssssh.: $p = 0,2639$

Tekn. konklusion: Signifikanssandsynligheden er større end 0,05 og dermed kan H_0 ikke forkastes.

Let konklusion: Vi kan ikke afvise hypotesen om, at den forventede afkastningsgrad er den samme i alle regnskabsår. M.a.o.: Vi kan ikke afvise, at den forventede afkastningsgrad ikke varierer fra regnskabsår til regnskabsår, hvilket stemmer overens med de optegnede boxplots i delspg. a).



OPGAVE 6

6A): Variablene i den reducerede regressionsmodel (ved et signifikansniveau på 5%) er *Region*, *Selskabsform*, *Nettoomsætning/balance*, *Anlægsaktiver/balance*, *Kortfr.gæld/balance*, *Status*.

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Region	5	5	3326.9981	3.2622	0.0062*
Selskabsform	1	1	3868.1278	18.9637	<.0001*
Nettoomsætning/balance	1	1	5973.2874	29.2843	<.0001*
Anlægsaktiver/balance	1	1	3967.8412	19.4525	<.0001*
Kortfr.gæld/balance	1	1	1902.1954	9.3256	0.0023*
Status	1	1	3642.4769	17.8574	<.0001*

6B): NB: Responsvariablen *Afkastningsgrad* er opgjort i procent (jf. opgaveteksten), og det er forklaringen på, at der i fortolkningen af parameterestimaterne nedenfor hele tiden angives "procent".

Parameterestimaterne hørende til variablen *Region* er...

- $\hat{\beta}_{\text{Region,Nordjylland}} = -2,4517$
- $\hat{\beta}_{\text{Region,Midtjylland}} = -3,3487$
- $\hat{\beta}_{\text{Region,Syddanmark}} = -4,6839$
- $\hat{\beta}_{\text{Region,Hovedstaden}} = -1,0912$
- $\hat{\beta}_{\text{Region,Sjælland}} = -3,3173$
- $\hat{\beta}_{\text{Region,Grønland}} = 0$

Resultatet af den estimerede model:

Indicator Function Parameterization						
Term	Estimate	Std Error	t Ratio	Prob> t	Lower 95%	Upper 95%
Intercept	10.472209	3.346523	3.13	0.0018*	3.9084438	17.035975
Region[Nordjylland]	-2.451736	3.200986	-0.77	0.4438	-8.730049	3.8265771
Region[Midtjylland]	-3.348676	2.921576	-1.15	0.2519	-9.078964	2.3816118
Region[Syddanmark]	-4.683878	2.926767	-1.60	0.1097	-10.42435	1.0565898
Region[Hovedstaden]	-1.091202	2.863969	-0.38	0.7032	-6.7085	4.5260971
Region[Sjælland]	-3.317274	2.945061	-1.13	0.2602	-9.093625	2.4590769
Selskabsform[A/S]	-3.13387	0.719647	-4.35	<.0001*	-4.545364	-1.722377
Nettoomsætning/balance	2.0219295	0.373636	5.41	<.0001*	1.2890921	2.7547668
Anlægsaktiver/balance	-6.774289	1.535945	-4.41	<.0001*	-9.786843	-3.761735
Kortfr.gæld/balance	-5.870195	1.922268	-3.05	0.0023*	-9.640472	-2.099918
Status[Aktiv]	4.2915333	1.015554	4.23	<.0001*	2.299657	6.2834097

Det betyder, at den forventede afkastningsgrad alt andet lige er...

- 2,4517 procent lavere, hvis virksomheden er beliggende i region Nordjylland, end hvis virksomheden er beliggende i Grønland
- 3,3487 procent lavere, hvis virksomheden er beliggende i region Midtjylland, end hvis virksomheden er beliggende i Grønland
- 4,6839 procent lavere, hvis virksomheden er beliggende i region Syddanmark, end hvis virksomheden er beliggende i Grønland
- 1,0912 procent lavere, hvis virksomheden er beliggende i region Hovedstaden, end hvis virksomheden er beliggende i Grønland
- 3,3173 procent lavere, hvis virksomheden er beliggende i region Sjælland, end hvis virksomheden er beliggende i Grønland

Parameterestimatet hørende til variablen...

- *Selskabsform* er $\hat{\beta}_{\text{Selskabsform,A/S}} = -3,1339$, hvis virksomheden er et aktieselskab, og $\hat{\beta}_{\text{Selskabsform,APS}} = 0$, hvis virksomheden er et anpartsselskab. Den forventede afkastningsgrad falder således med 3,1339 procent, såfremt virksomheden er et aktie- og ikke et anpartsselskab.
- *Nettoomsætning/balance* er $\hat{\beta}_{\text{Nettoomsætning/balance}} = 2,0219$, således at den forventede afkastningsgrad stiger med $2,0219 \cdot 0,01 = 0,020219$ procent, hvis virksomhedens omsætning relativt til den samlede balance stiger med 0,01.
- *Anlægsaktiver/balance* er $\hat{\beta}_{\text{Anlægsaktiver/balance}} = -6,7743$, således at den forventede afkastningsgrad falder med $6,7743 \cdot 0,01 = 0,067743$ procent, hvis virksomhedens anlægsaktiver relativt til den samlede balance stiger med 0,01.



- **Kortfr.gæld/balance** er $\hat{\beta}_{\text{Kortfr.gæld/balance}} = -5,8702$, således at den forventede afkastningsgrad falder med $5,8702 \cdot 0,01 = 0,058702$ procent, hvis virksomhedens kortfristede gæld relativt til den samlede balance stiger med 0,01.
- **Status** er $\hat{\beta}_{\text{Status,aktiv}} = 4,2915$, hvis virksomheden er aktiv, og $\hat{\beta}_{\text{Status,Konkurs}} = 0$, hvis virksomheden er gået konkurs i løbet af indeværende år. Den forventede afkastningsgrad stiger således med 4,2915 procent, såfremt virksomheden er aktiv, dvs. såfremt virksomheden ikke er gået konkurs i løbet af indeværende år.

6c): R^2 -værdien for den oprindelige model er 0,0892, mens den for den reducerede model er 0,0770. R^2 -værdien har den egenskab at den altid stiger, desto flere variable/parametre regressionsmodellen indeholder, hvilket er forklaringen på, at den oprindelige model har den højeste R^2 -værdi.

Summary of Fit	
RSquare	0,077021
RSquare Adj	0,071559
Root Mean Square Error	14,282
Mean of Response	5,588713
Observations (or Sum Wgts)	1701

6d): Et 95%-konfidensinterval for betydningen af selskabsform er $[-4,5454; -1,7224]$. Med 95% sandsynlighed vil den forventede afkastningsgrad dermed falde med mellem 4,5454 og 1,7224 procent, såfremt virksomheden er et aktie- og ikke et anpartsselskab.

6e): Den forventede afkastningsgrad...

- stiger med $2,0219 \cdot 0,01 = 0,020219$ procent, såfremt **Nettoomsætning/balance** stiger med 0,01
- stiger med $5,8702 \cdot 0,01 = 0,058702$ procent, såfremt **Kortfr.gæld/balance** falder med 0,01

Det betyder dermed alt andet lige mere for den forventede afkastningsgrad, såfremt den kortfristede gæld falder, end såfremt nettoomsætningen stiger.

6f): Den forventede afkastningsgrad...

- stiger med $\hat{\beta}_{\text{Region,Hovedstaden}} - \hat{\beta}_{\text{Region,Syddanmark}} = -1,0912 - (-4,6839) = 3,5927$ procent, såfremt virksomheden er beliggende i region Hovedstaden frem for i region Syddanmark.
- stiger med $\hat{\beta}_{\text{Region,Hovedstaden}} - \hat{\beta}_{\text{Region,Grønland}} = -1,0912 - 0 = -1,0912$ procent (dvs. falder med 1,0912 procent), såfremt virksomheden er beliggende i region Hovedstaden frem for i Grønland.



OPGAVE 7

7A): Krydstabel mellem virksomhedens status og selskabsform:

		Status		
		Aktiv	Konkurs	Total
Selskabsform	Count			
	Total %			
	Col %			
	Row %			
	A/S	762	85	847
		44,80	5,00	49,79
		52,26	34,98	
APS		89,96	10,04	
		696	158	854
		40,92	9,29	50,21
		47,74	65,02	
Total		81,50	18,50	
		1458	243	1701
		85,71	14,29	

Sandsynligheden for at en tilfældigt valgt virksomhed er et anpartsselskab, når vi udelukkende ser på virksomheder, der er gået konkurs, er:

$$\begin{aligned}
 & P(\{\text{virksomhed er APS}\} | \{\text{virksomhed er gået konkurs}\}) \\
 &= \frac{P(\{\text{virksomhed er APS}\} \cap \{\text{virksomhed er gået konkurs}\})}{P(\{\text{virksomhed er gået konkurs}\})} \\
 &= \frac{\frac{158}{1.701}}{\frac{243}{1.701}} = 65,02\%
 \end{aligned}$$

Sandsynligheden for at en tilfældigt valgt virksomhed er et anpartsselskab og er gået konkurs er:

$$P(\{\text{virksomhed er APS}\} \cap \{\text{virksomhed er gået konkurs}\}) = \frac{158}{1.701} = 9,29\%$$



7b): Andelen af konkursramte virksomheder indenfor hver af de to selskabsformer er

$$\hat{p}_{A/S} = \frac{85}{847} = 10,04\% \quad \hat{p}_{APS} = \frac{158}{854} = 18,50\%$$

Test af nulhypotese om at den forventede andel af konkursramte virksomheder er den samme blandt aktie- og anpartsselskaber:

Contingency Table				
Selskabsform	Status			Total
	Aktiv	Konkurs	Total	
Count				
Total %				
Col %				
Row %				
Expected				
A/S	762	85	847	
	44,80	5,00	49,79	
	52,26	34,98		
	89,96	10,04		
APS	726	121	847	
	696	158	854	
	40,92	9,29	50,21	
	47,74	65,02		
	81,50	18,50		
	732	122		
Total	1458	243	1701	
	85,71	14,29		

Two Sample Test for Proportions			
Description	Proportion Difference	Lower 95%	Upper 95%
P(Konkurs A/S)-P(Konkurs APS)	-0,08466	-0,11748	-0,05143
Adjusted Wald Test		Prob	
P(Konkurs A/S)-P(Konkurs APS) ≥ 0		1,0000	
P(Konkurs A/S)-P(Konkurs APS) ≤ 0		<,0001*	
P(Konkurs A/S)-P(Konkurs APS) = 0		<,0001*	

Antagelser: Observationerne af om en virksomhed er gået konkurs eller ej er indbyrdes uafhængige. Endvidere er det forventede antal observationer i hver af krydstabellens celler større end 5. Sidstnævnte antagelse er opfyldt, da alle celler har et forventet antal observationer på 121 eller mere.

Hypoteser: $H_0 : p_{A/S} = p_{APS}$

$H_a : p_{A/S} \neq p_{APS}$

Signifikansssh.: $p < 0,0001$

Tekn. konklusion: Signifikanssandsynligheden er mindre end 0,05 og dermed forkastes H_0 .

Let konklusion: Der er ikke belæg for en påstand om, at andelen af konkursramte virksomheder er den samme blandt aktie- og anpartsselskaber. Konklusionen er ikke overraskende, eftersom andelen af konkursramte aktieselskaber estimeres til 10,04%, mens den tilsvarende andel blandt anpartsselskaber estimeres til 18,50%. Der er således markant flere konkurser blandt anparts- end blandt aktieselskaber.



7c): Datamaterialet indeholder $n_{A/S} = 847$ observationer af aktieselskaber, hvoraf en andel på $\hat{p}_{APS} = 10,04\% (= 0,1004)$ er gået konkurs, og tilsvarende $n_{APS} = 854$ observationer af anpartsselskaber, hvoraf en andel på $\hat{p}_{A/S} = 18,50\% (= 0,1850)$ er gået konkurs. Det giver et 95%-konfidensinterval for forskellen på de to forventede værdier på

$$\begin{aligned} & \left[\hat{p}_{A/S} - \hat{p}_{APS} - z_{0,025} \cdot \sqrt{\frac{\hat{p}_{A/S}(1 - \hat{p}_{A/S})}{n_{A/S}} + \frac{\hat{p}_{APS}(1 - \hat{p}_{APS})}{n_{APS}}}; \hat{p}_{A/S} - \hat{p}_{APS} + z_{0,025} \cdot \sqrt{\frac{\hat{p}_{A/S}(1 - \hat{p}_{A/S})}{n_{A/S}} + \frac{\hat{p}_{APS}(1 - \hat{p}_{APS})}{n_{APS}}} \right] \\ &= \left[0,1004 - 0,1850 - 1,96 \cdot \sqrt{\frac{0,1004 \cdot (1 - 0,1004)}{847} + \frac{0,1850 \cdot (1 - 0,1850)}{854}}; \right. \\ & \quad \left. 0,1004 - 0,1850 + 1,96 \cdot \sqrt{\frac{0,1004 \cdot (1 - 0,1004)}{847} + \frac{0,1850 \cdot (1 - 0,1850)}{854}} \right] \\ &= [-0,1175; -0,0514] \end{aligned}$$

Med 95% sandsynlighed vil forskellen mellem andelen af konkursramte virksomheder blandt hhv. aktie- og anpartsselskaber dermed ligge mellem -0,1175 og -0,0514. M.a.o.: Andelen af konkursramte aktieselskaber vil med 95% sandsynlighed ligge mellem 5,14% og 11,75% lavere end den tilsvarende andel af konkursramte anpartsselskaber.

Desto flere observationer datamaterialet indeholder, desto mere præcist kan vi estimere de ukendte andele af konkursramte virksomheder, og desto smallere vil konfidensintervallet dermed blive, alt andet lige.



OPGAVE 8

8A): Krydstabel mellem virksomhedens status og hvilken region, virksomheden er beliggende i:

	Count	Aktiv	Konkurs	Total
Total %				
Col %				
Row %				
Nordjylland	71	17	88	
	4,17	1,00	5,17	
	4,87	7,00		
	80,68	19,32		
Midtjylland	293	41	334	
	17,23	2,41	19,64	
	20,10	16,87		
	87,72	12,28		
Syddanmark	246	54	300	
	14,46	3,17	17,64	
	16,87	22,22		
	82,00	18,00		
Hovedstaden	608	77	685	
	35,74	4,53	40,27	
	41,70	31,69		
	88,76	11,24		
Sjælland	218	50	268	
	12,82	2,94	15,76	
	14,95	20,58		
	81,34	18,66		
Grønland	22	4	26	
	1,29	0,24	1,53	
	1,51	1,65		
	84,62	15,38		
Total	1458	243	1701	
	85,71	14,29		

Andelen af konkursramte virksomheder i hver af de 6 regioner er:

$$\hat{p}_{Nordjyl} = 19,32\% \quad \hat{p}_{Midtjyl} = 12,28\% \quad \hat{p}_{SydDK} = 18,00\% \quad \hat{p}_{Hovedst} = 11,24\% \quad \hat{p}_{Sjl} = 18,66\% \quad \hat{p}_{Grn} = 15,38\%$$

Region Nordjylland er med en andel på 19,32% den af de 6 regioner med størst andel konkursramte virksomheder.

8B): Sandsynligheden for at en tilfældigt valgt virksomhed er konkursramt, når vi udelukkende ser på virksomheder beliggende i Nord- eller Midtjylland, er:

$$\begin{aligned}
 & P(\{\text{virksomhed er konkurs}\} | \{\text{virksomhed ligger i Nord- eller Midtjylland}\}) \\
 &= \frac{P(\{\text{virksomhed er konkurs}\} \cap \{\text{virksomhed ligger i Nord- eller Midtjylland}\})}{P(\{\text{virksomhed ligger i Nord- eller Midtjylland}\})} \\
 &= \frac{\frac{17+41}{1.701}}{\frac{88+334}{1.701}} = 13,74\%
 \end{aligned}$$



8c): Test af nulhypotese om at der er uafhængighed mellem status og region:

Antagelser:	Observationerne af om en virksomhed er gået konkurs eller ej er indbyrdes uafhængige. Endvidere er det forventede antal observationer i hver af krydstabellens celler større end 5. Sidstnævnte antagelse er næsten opfyldt, da alle celler med undtagelse af én har et forventet antal observationer på 12 eller mere. Kun cellen med konkurser i Grønland opfylder ikke betingelsen om et forventet antal større end 5.
Hypoteser:	H_0 : Uafhængighed mellem Status og Region H_a : Afhængighed mellem Status og Region
Signifikanssssh.:	$p = 0,0078$
Tekn. konklusion:	Signifikanssandsynligheden er på 0,0078 og dermed mindre end 0,05 og dermed forkastes H_0 .
Let konklusion:	Der er ikke belæg for en påstand om, at der er uafhængighed mellem om en virksomhed er gået konkurs og hvilken region, virksomheden er beliggende i. M.a.o. er der belæg for en påstand om, at der er en sammenhæng (afhængighed) mellem sandsynligheden for at en virksomhed går konkurs, og dens geografiske placering.

Contingency Table				
		Status		
		Aktiv	Konkurs	Total
Count				
Total %				
Col %				
Row %				
Expected				
Region	Midtjylland	80,68	19,32	
		75,4286	12,5714	
		293	41	334
		17,23	2,41	19,64
		20,10	16,87	
Syddanmark		87,72	12,28	
		286,286	47,7143	
		246	54	300
		14,46	3,17	17,64
		16,87	22,22	
Hovedstaden		82,00	18,00	
		257,143	42,8571	
		608	77	685
		35,74	4,53	40,27
		41,70	31,69	
Sjælland		88,76	11,24	
		587,143	97,8571	
		218	50	268
		12,82	2,94	15,76
		14,95	20,58	
Grønland		81,34	18,66	
		229,714	38,2857	
		22	4	26
		1,29	0,24	1,53
		1,51	1,65	
Total		84,62	15,38	
		22,2857	3,71429	
Total		1458	243	1701
		85,71	14,29	

Tests			
	N	DF	-LogLike RSquare (U)
	1701	5	7,7066135 0,0110
Test	ChiSquare	Prob>ChiSq	
Likelihood Ratio	15,413	0,0087*	
Pearson	15,696	0,0078*	

Af krydstabellen fremgår det, at der er tale om en "Udkantsdanmark"-effekt, idet de klart lavest konkursandele findes i nærheden af landets to største byer, Århus (region Midtjylland) og København (region Hovedstaden). Eftersom at der er særligt mange virksomheder hjemmehørende i region Hovedstaden (685 ud af datamaterialets i alt 1.701 virksomheder) bliver den lavere konkursandel i region Hovedstaden (og delvist også den i region Midtjylland) udslagsgivende for, at nulhypotesen forkastes.



8d): Test af nulhypotese om at der er uafhængighed mellem status og region (eksklusiv region Hovedstaden):

Antagelser:

Observationerne af om en virksomhed er gået konkurs eller ej er indbyrdes uafhængige. Endvidere er det forventede antal observationer i hver af krydstabellens celler større end 5. Sidstnævnte antagelse er næsten opfyldt, da alle celler med undtagelse af én har et forventet antal observationer på 14 eller mere. Kun cellen med konkurser i Grønland opfylder ikke betingelsen om et forventet antal større end 5.

Hypoteser:

H_0 : Uafhængighed mellem Status og Region

H_a : Afhængighed mellem Status og Region

Signifikanssssh.:

$p = 0,1790$

Tekn. konklusion:

Signifikanssandsynligheden er på 0,1790 og dermed større end 0,05 og dermed kan H_0 ikke forkastes.

Let konklusion:

Vi kan ikke afvise hypotesen om, at andelen af konkursramte selskaber er den samme i alle regioner med undtagelse af region Hovedstaden. Selv om der er en lavere andel i region Midtjylland end i de øvrige regioner (jf. delspg. c), er det ikke nok til at afvise hypotesen.

Contingency Analysis of Status By Re

Contingency Table

	Status		
	Aktiv	Konkurs	Total
Count			
Total %			
Col %			
Row %			
Expected			
Nordjylland	71	17	88
	6,99	1,67	8,66
	8,35	10,24	
	80,68	19,32	
	73,622	14,378	
Midtjylland	293	41	334
	28,84	4,04	32,87
	34,47	24,70	
	87,72	12,28	
	279,429	54,5709	
Syddanmark	246	54	300
	24,21	5,31	29,53
	28,94	32,53	
	82,00	18,00	
	250,984	49,0157	
Sjælland	218	50	268
	21,46	4,92	26,38
	25,65	30,12	
	81,34	18,66	
	224,213	43,7874	
Grønland	22	4	26
	2,17	0,39	2,56
	2,59	2,41	
	84,62	15,38	
	21,752	4,24803	
Total	850	166	1016
	83,66	16,34	

Tests

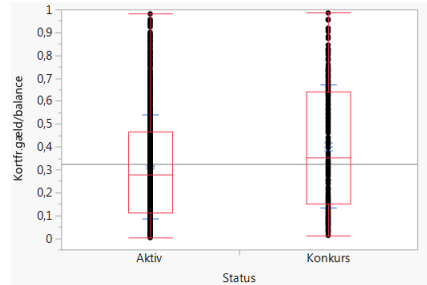
	N	DF	-LogLike	RSquare (U)
	1016	4	3,2571308	0,0072

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	6,514	0,1639
Pearson	6,282	0,1790



OPGAVE 9

9A): Boxplot over fordelingen af afkastningsgraden for henholdsvis aktive og konkursramte virksomheder.



Figuren viser, at niveauet af kortfristet gæld², som angivet ved medianen, er lidt lavere (omkring 0,28) for aktive end for konkursramte virksomheder (omkring 0,35). Der er således umiddelbart tegn på, at virksomheder der går konkurs har mere kortfristet gæld end øvrige virksomheder. Variationen i afkastningsgrad, som angivet ved interkvartilbredden, er ligeledes mindre for aktive end for konkursramte virksomheder. Der er således også tegn på mindre variation i størrelsen af kortfristet gæld blandt aktive end blandt konkursramte virksomheder.

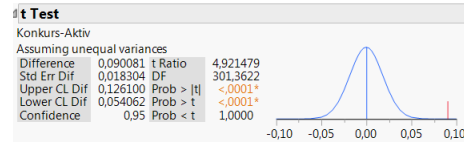
9B): Middelværdi og standardafvigelse for fordelingen af kortfristet gæld for henholdsvis aktive og konkursramte virksomheder understøtter tendenserne i boxplottene beskrevet i delspg. a). En lavere middelværdi for aktive virksomheder (0,3122 mod de konkursramtes 0,4022) indikerer mindre kortfristet gæld i aktive virksomheder, mens en lavere standardafvigelse (0,2264 mod de konkursramtes 0,2699) indikerer mindre variation i niveauet af kortfristet gæld blandt aktive virksomheder.

Means and Std Deviations						
Level	Number	Mean	Std Dev	Std Err	Lower 95%	Upper 95%
Aktiv	1458	0,312156	0,226369	0,00593	0,30053	0,32378
Konkurs	243	0,402237	0,269945	0,01732	0,36813	0,43635

²I besvarelsen af opgave 9 underforstås, at alle tal vedrørende kortfristet gæld udtrykker niveauet af kortfristet gæld relativt til den samlede balance.



9c): Test af nulhypotese om at det forventede niveau af kortfristet gæld er det samme for aktive og konkursramte virksomheder:



- Antagelser:** Observationerne af de enkelte virksomheders kortfristede gæld er indbyrdes uafhængige og approksimativt normalfordelte; normalfordelte indenfor hver gruppe af virksomheder (aktive hhv. konkursramte).
- Hypoteser:** $H_0 : \mu_{konkurs} = \mu_{aktiv}$
 $H_a : \mu_{konkurs} \neq \mu_{aktiv}$
- Signifikansssh.:** $p < 0,0001$
- Tekn. konklusion:** Signifikanssandsynligheden er mindre end 0,05 og dermed forkastes H_0 .
- Let konklusion:** Der er ikke belæg for en påstand om, at det forventede niveau af kortfristet gæld er det samme for aktive og konkursramte virksomheder. På baggrund af resultaterne i delspg. a) og b) er den konklusion ikke overraskende, eftersom det forventede niveau af kortfristet gæld estimeres til 0,3122 for aktive og 0,4022 for konkursramte virksomheder. Der er således væsentligt mere kortfristet gæld blandt konkursramte end blandt aktive virksomheder. Det kan være et tegn på, at kortfristet gæld er en af de faktorer der (næppe særligt overraskende) er udslagsgivende for om en virksomhed går konkurs eller ej.



OPGAVE 10

Spørgsmålet om der er sammenhæng mellem afkastningsgraden på aktive og konkursramte virksomheder kan løses meget nemt og hurtigt ved blot at referere til resultatet i opgave 6 b), som viste, at der var en signifikant lavere forventet afkastningsgrad for konkursramte virksomheder (4,2915 procent point lavere).

Alternativt kan opgaven eksempelvis løses ved at teste for, om den forventede afkastningsgrad er den samme for aktive og konkursramte virksomheder.

10A): Test af nulhypotese om at den forventede afkastningsgrad er den samme for aktive og konkursramte virksomheder:



Antagelser: Observationerne af de enkelte virksomheders afkastningsgrad er indbyrdes uafhængige og approksimativt normalfordelte; normalfordelte indenfor hver gruppe af virksomheder (aktive hhv. konkursramte).

Hypoteser: $H_0 : \mu_{konkurs} = \mu_{aktiv}$

$H_a : \mu_{konkurs} \neq \mu_{aktiv}$

Signifikansssh.: $p = 0,0004$

Tekn. konklusion: Signifikanssandsynligheden er mindre end 0,05 og dermed forkastes H_0 .

Let konklusion: Der er ikke belæg for en påstand om, at den forventede afkastningsgrad er den samme blandt aktive og konkursramte virksomheder. I det omfang at afkastningsgrad udtrykker virksomhedens evne til at tjene penge er det næppe en overraskende konklusion, idet aktive virksomheder dermed - alt andet lige - er bedre til at tjene penge end de virksomheder, der er gået konkurs. Det indikerer, at virksomhedens indtjening (igen, ikke overraskende) har afgørende betydning for, om en virksomhed går konkurs eller ej.