

Copenhagen Business School  
Dataanalyse, HD 1. del

**EKSAMENSOPGAVE**  
**ORDINÆR PRØVE, EFTERÅR 2017**

4 timers skriftlig eksamen

Opgaven består af i alt **31** delspørgsmål fordelt på **9** opgaver. Ved bedømmelsen vægtes alle opgavens delspørgsmål ligeligt med undtagelse af delspørgsmål 9 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.



## ØLSALG I DANMARK

### BAGGRUND OG DATABESKRIVELSE

Salget af øl på det danske marked er præget af stor konkurrence, både mellem forskellige produkter og mellem forskellige butikskæder. I denne opgave vil vi se på, hvorledes det er muligt at bruge historisk salgsdata til at få et mere præcist billede af de vigtigste markedstendenser indenfor detailsalget af øl i Danmark.

Datamaterialet i JMP-filen *ØlsalgDK.jmp* indeholder 3 års ugentlige observationer af salget af 33 cl dåseøl af fire forskellige mærker (Grøn Tuborg, Carlsberg Pilsner, Heineken Premium og Royal Export) i to forskellige pakkestørrelser: 6 stk ("six pack") og 24 stk ("en ramme") i syv forskellige supermarkeds-kæder (Bilka, Føtex, Discount<sup>1</sup>, Købmandshuset, Kvickly, Dagli/Lokal/Irma<sup>2</sup> og SuperBrugsen).

Datamaterialet indeholder information om den omsatte mængde (dvs. én six pack = 2 liter og én ramme = 8 liter), om den gennemsnitlige ugentlige salgspris (angivet i kr. pr. liter) samt information om i hvilket omfang, salget er sket til en midlertidigt nedsat pris ("rabatpris"). Hvis mindst halvdelen af ugens omsatte mængde (målt i liter) er solgt til rabatpris, betegnes ugen som en "kampagneuge", dvs. en uge hvor der har været ekstraordinært omfattende prisreduktion i forhold til normalprisen.<sup>3</sup>

Eksamensopgaven skal besvares på baggrund af datamaterialet i *ØlsalgDK.jmp*, hvis indhold er nærmere beskrevet i tabellen på næste side.

---

<sup>1</sup>'Discount' indeholder det samlede salg i Netto, Fakta, Rema og Kiwi butikkerne.

<sup>2</sup>"Daglig/Lokal/Irma" indeholder det samlede salg i Dagli'Brugsen, LokalBrugsen og Irma butikkerne.

<sup>3</sup>Hvorvidt en given uge er en kampagneuge eller ej for et givet ølmærke, afhænger både af hvilken pakkestørrelse og hvilken supermarkeds-kæde omsætningstallene stammer fra. Det kan således sagtens være, at en given uge er kampagneuge for Royal Export 6 stk. pakker i SuperBrugsen, mens den samme uge ikke er kampagneuge for Royal Export 24 stk. pakker i SuperBrugsen (fordi der i så fald kun er tilbud i SuperBrugsen på Royal Export i 6 stk. pakker og ikke i 24 stk. pakker i den pågældende uge).



VARIABELNAVN	FORKLARING
<i>Generelt</i>	
<i>Dato</i>	Sidste dag i den betragtede uges omsætning
<i>Kæde</i>	Hvilken supermarkedskæde omsætningstallene stammer fra (med de mulige værdier "Bilka", "Føtex", "Discount", "Købmandshuset", "Kvickly", "Dagli/Lokal/Irma", "SuperBrugsen")
<i>Pakkestørrelse</i>	Antal 33 cl dåser i den solgte vare (med de mulige værdier "6 stk", "24 stk")
<i>Prisinformation (angivet i kr. pr. liter)</i>	
<i>PrisCarlsberg</i>	Pris på Carlsberg Pilsner
<i>PrisTuborg</i>	Pris på Grøn Tuborg
<i>PrisHeineken</i>	Pris på Heineken Premium
<i>PrisRoyal</i>	Pris på Royal Export
<i>Omsætningsinformation (angivet i 100 liter)</i>	
<i>VolumeCarlsberg</i>	Ugentligt omsætning af Carlsberg Pilsner
<i>VolumeTuborg</i>	Ugentligt omsætning af Grøn Tuborg
<i>VolumeHeineken</i>	Ugentligt omsætning af Heineken Premium
<i>VolumeRoyal</i>	Ugentligt omsætning af Royal Export
<i>Reklamekampagneinformation I (angivet i procent)</i>	
<i>ShPromoCarlsberg</i>	Andel af den ugentlige omsætning (i liter) af Carlsberg Pilsner, der er solgt til rabatpris
<i>ShPromoTuborg</i>	Andel af den ugentlige omsætning (i liter) af Grøn Tuborg, der er solgt til rabatpris
<i>ShPromoHeineken</i>	Andel af den ugentlige omsætning (i liter) af Heineken Premium, der er solgt til rabatpris
<i>ShPromoRoyal</i>	Andel af den ugentlige omsætning (i liter) af Royal Export, der er solgt til rabatpris
<i>Reklamekampagneinformation II</i>	
<i>UgeCarlsberg</i>	Hvorvidt den betragtede uge er en kampagneuge for Carlsberg Pilsner (med de mulige værdier "Almindelig", "Kampagne")
<i>UgeTuborg</i>	Hvorvidt den betragtede uge er en kampagneuge for Grøn Tuborg (med de mulige værdier "Almindelig", "Kampagne")
<i>UgeHeineken</i>	Hvorvidt den betragtede uge er en kampagneuge for Heineken Premium (med de mulige værdier "Almindelig", "Kampagne")
<i>UgeRoyal</i>	Hvorvidt den betragtede uge er en kampagneuge for Royal Export (med de mulige værdier "Almindelig", "Kampagne")



### OPGAVE 1 (VOLUMEN)

For at få et overblik over hvor meget øl af mærket Royal Export der sælges på det danske marked, ser vi på den ugentligt omsatte mængde af Royal Export, som angivet i variabelen *VolumeRoyal*, for hver af de to pakkestørrelser, som angivet i variabelen *Pakkestørrelse*.

- a). **Tegn** for hver af de to pakkestørrelser et histogram og et boxplot over fordelingen af variabelen *VolumeRoyal*.  
**Giv** en kort beskrivelse af de to fordelinger af variabelen *VolumeRoyal* på baggrund af histogram, boxplot og relevante nøgletal for centrum og variation.
- b). **Angiv og fortolk** medianen i fordelingen af variabelen *VolumeRoyal* for 6 stk. pakker.  
**Gør rede for** om median eller middelværdi (eller begge) er retvisende mål for centrum i fordelingen af variabelen *VolumeRoyal* for 6 stk. pakker.

I de resterende delspørgsmål i denne opgave ser vi udelukkende på omsætningen af Royal Export i 6 stk. pakker i SuperBrugsen.

*Vink: Sæt f.eks. begge de to variable *Kæde* og *Pakkestørrelse* under "By" i JMPs "Distribution"-funktion.*

- c). **Tegn** den kumulerede fordelingsfunktion (CDF) for variabelen *VolumeRoyal*.  
**Gør rede for** hvorvidt den kumulerede fordelingsfunktion indikerer, at der (i gennemsnit) mindst hver fjerde uge sælges mere end 7.000 liter Royal Export i 6 stk. pakker i SuperBrugsen.
- d). **Beregn** et 95%-konfidensinterval for den forventede omsatte mængde af Royal Export i 6 stk. pakker i SuperBrugsen.  
**Gør rede for** hvorledes konfidensintervallet ændres, såfremt standardafvigelsen af variabelen *VolumeRoyal* stiger.
- e). **Gør rede for** om der er statistisk belæg for en nulhypotese om, at den forventede ugentlige omsatte mængde af Royal Export i 6 stk. pakker i SuperBrugsen er 5.250 liter. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

### OPGAVE 2 (VOLUMEN OG PRIS HHV. REKLAME)

Hvor meget Royal Export øl der sælges i en given uge hænger blandt andet sammen med hvilken pris, produktet sælges til. Derudover er det naturligt at forvente, at salget også påvirkes af, i hvilket omfang der gennemføres tilbudskampagner for produktet samt af salget af konkurrerende produkter.

- a). **Tegn** scatterplots af variablene *VolumeRoyal*, *PrisRoyal* og *ShPromoRoyal* mod hinanden.  
**Beregn** variablenes indbyrdes korrelation og **gør rede for** mulig intuition bag fortegnene på de beregnede korrelationer.



- b). **Tegn** scatterplots af variablene *VolumeRoyal*, *VolumeHeineken* og *ShPromoHeineken* mod hinanden.  
**Beregn** variablenes indbyrdes korrelation og **gør rede for** mulig intuition bag fortegnene på de beregnede korrelationer.
- c). **Gør rede for** i hvilket omfang korrelationerne i delspg. b) kan siges at udtrykke en kausal sammenhæng.

### OPGAVE 3 (PRIS OG PAKKESTØRRELSE)

Royal Export sælges i forskellige pakkestørrelser og et oplagt spørgsmål er derfor, hvorvidt der er forskel på prisen pr. liter afhængig af hvilken pakkestørrelse, man køber.

- a). **Tegn** for hver af de to pakkestørrelser et histogram over fordelingen af *PrisRoyal* og en tilhørende normalfordelingskurve.  
**Gør** for hver af de to pakkestørrelser **rede for** (på baggrund af figurene) om normalfordelingen ser ud til at give en rimelig beskrivelse af fordelingen af prisen pr. liter Royal Export.
- b). **Gør rede for** om der er statistisk belæg for en nulhypotese om, at den forventede pris pr. liter Royal Export er den samme for pakker med henholdsvis 6 og 24 stk. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.
- c). **Beregn** et 95%-konfidensinterval for forskellen i den forventede pris pr. liter Royal Export for pakker med henholdsvis 6 og 24 stk. og **giv** en fortolkning af intervallet.

### OPGAVE 4 (PRIS OG KÆDE)

Prisen på Royal Export varierer ikke kun som følge af reklamekampagner og påvirkning fra konkurrerende produkter. Royal Export sælges i en række forskellige supermarkeds kæder, og ikke altid til samme pris i de forskellige kæder. Det er derfor relevant at undersøge, om der er systematiske forskelle i prisen på tværs af de forskellige kæder.

- a). **Tegn** et boxplot over fordelingen af variablen *PrisRoyal* for hver værdi af variablen *Kæde*.  
**Beregn** middelværdi og standardafvigelse af variablen *PrisRoyal* for hver værdi af variablen *Kæde*.
- b). **Gør rede for** at der ikke er statistisk belæg for en nulhypotese om, at den forventede pris pr. liter Royal Export er den samme i alle kæder. **Angiv** i den forbindelse antagelser, hypoteser og signifikanssandsynlighed.  
*Vink: Gennemfør et passende valgt statistisk test og konkluder på baggrund heraf, at den relevante nulhypotese forkastes.*



- c). **Gør rede for** at der er statistisk belæg for en nulhypotese om, at den forventede pris pr. liter Royal Export er den samme i mindst tre af de betragtede kæder. **Angiv** i den forbindelse antagelser, hypoteser og signifikanssandsynlighed.

*Vink: Vælg tre kæder hvis middelværdier jf. delspg. a) ligger tæt på hinanden og udfør et statistisk test kun med data for disse tre kæder.*

#### OPGAVE 5 (VOLUMEN, SAMLET ANALYSE)

Mange faktorer påvirker salget af Royal Export, og der er således behov for en samlet vurdering af de enkelte faktorerers påvirkning. Vi ser derfor på en lineær regressionsmodel med variabelen *VolumeRoyal* som responsvariabel og variablene *PrisRoyal*, *Kæde*, *Pakkestørrelse*, *VolumeCarlsberg*, *VolumeTuborg*, *VolumeHeineken*, *ShPromoCarlsberg*, *ShPromoTuborg*, *ShPromoHeineken*, *ShPromoRoyal* som forklarende variable.

- a). **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. Som signifikansniveau anvendes som sædvanlig 5%.*
- b). **Gør rede for** om den reducerede regressionsmodel ville indeholde flere eller færre forklarende variable, hvis vi ændrede signifikansniveauet fra det sædvanlige 5% til 1% og **forklar** hvorfor.
- c). **Angiv** estimater for regressionskoefficienter hørende til variablene i den reducerede model fra delspg. a) og **giv** en fortolkning af estimaterne.
- d). **Beregn** et 95%-konfidensinterval for betydningen af en prisstigning på 1 kr. pr. liter Royal Export og **giv** en fortolkning af intervallet.
- e). **Beregn** et 95%-konfidensinterval for betydningen af en prisstigning på 0,50 kr. pr. liter Royal Export.
- f). **Beregn** på baggrund af den estimerede reducerede regressionsmodel det forventede salg (i liter) af Royal Export 6 stk-pakker i Føtex i en given uge, forudsat at regressionsmodellens øvrige forklarende variable sættes lig deres respektive middelværdier (beregnet på baggrund af hele datasættet).  
*Vink: Ved brug af "Profiler" i JMPs "Fit Model" regressionsoutput sættes alle numeriske forklarende variable pr. default lig deres respektive middelværdier beregnet på baggrund af hele datasættet.*
- g). **Gør rede for** hvordan det forventede salg i delspg. f) ændres, såfremt der i stedet er tale om salget af 24 stk.-pakker.  
**Gør rede for** hvordan den forventede salg i delspg. f) ændres, såfremt der i stedet er tale om salget i SuperBrugsen.



### OPGAVE 6 (VOLUMEN OG REKLAME)

Reklamekampagner er i mange sammenhænge et yderst effektivt middel til at øge salget af et givet produkt. For at få en mere detaljeret forståelse af, præcis hvor meget reklamekampagner<sup>4</sup> påvirker salget af Royal Export, ser vi på fordelingen af salget af Royal Export (variablen *VolumeRoyal*) i 24 stk-pakker i discountbutikker (*Pakkestørrelse*="24 stk", *Kæde*="Discount") i henholdsvis almindelige uger og kampagneuger, som angivet i variablen *UgeRoyal*.

- Tegn** et histogram over fordelingen af salget af Royal Export i 24 stk-pakker i discountbutikker i en kampagneuge. *Vink: Sæt f.eks. alle tre variable Kæde, Pakkestørrelse, UgeRoyal under "By" i JMPs "Distribution"-funktion.*  
**Estimér** middelværdi og standardafvigelse i den normalfordeling, der bedst beskriver salget af Royal Export i 24 stk.-pakker i discountbutikker i en kampagneuge.
- Beregn** på baggrund af den i delspg. a) beregnede normalfordeling sandsynligheden for, at der sælges mindst 20.000 liter Royal Export i 24 stk-pakker i discountbutikker i løbet af en kampagneuge.  
**Beregn** på baggrund af den i delspg. a) beregnede normalfordeling sandsynligheden for, at der sælges mellem 20.000 og 70.000 liter Royal Export i 24 stk-pakker i discountbutikker i løbet af en kampagneuge.
- Beregn** på baggrund af en passende valgt normalfordeling sandsynligheden for, at der sælges mellem 20.000 og 70.000 liter Royal Export i 24 stk-pakker i discountbutikker i løbet af en almindelig uge.

### OPGAVE 7 (REKLAME OG KÆDE)

De forskellige supermarkeds kæder henvender sig til delvist forskellige kundesegmenter, og som følge heraf er det naturligt at forvente, at der bl.a. er forskel i deres valg af markedsføringsstrategi. Det er derfor interessant at undersøge, om der reelt er forskel på, hvor ofte Royal Export er på tilbud (dvs. sælges til rabatpris) i de forskellige kæder.

- Opstil** en krydstabel for de to variable *Kæde* og *UgeRoyal*.  
**Gør rede for** hvilken af de betragtede supermarkeds kæder, der i gennemsnit har færrest kampagneuger på et år.
- Gør rede for** om der er statistisk belæg for en nulhypotese om, at der er uafhængighed mellem hvilken kæde, omsætningen stammer fra, og om omsætningen stammer fra en kampagneuge eller ej. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

---

<sup>4</sup>I hele denne eksamensopgave opfattes "reklamekampagne" som synonymt med en midlertidig reduktion i den normale salgspris.



- c). **Gør rede for** om konklusionen i delspg. b) ændres, såfremt vi ser bort fra omsætningen i kæderne “Discount” og “Daglig/Lokal/Irma”. **Angiv** i den forbindelse antagelser, hypoteser, signifikanssandsynlighed samt teknisk og let forståelig konklusion.

#### OPGAVE 8 (REKLAME OG PAKKESTØRRELSE)

Ligesom de forskellige supermarkeds kæder har forskellige kundesegmenter, kan man overveje, om også de to forskellige pakkestørrelser henvender sig til forskellige segmenter. Hvis det er tilfældet, kan det være relevant med forskellige markedsføringsstrategier for de to pakkestørrelser (uafhængigt af i hvilken kæde de sælges).

- a). **Opstil** en krydstabel for de to variable *Pakkestørrelse* og *UgeRoyal*.  
**Estimer** andelen af kampagneuger for Royal Export for hver af de to pakkestørrelser.
- b). **Beregn** et 95%-konfidensinterval for andelen af kampagneuger på et år for Royal Export i 24-stk pakker. **Angiv** antagelser bag konfidensintervallet og **giv** en fortolkning af intervallet.
- c). **Gør rede for** om der er statistisk belæg for en nulhypotese om, at der ikke er forskel på andelen af kampagneuger for de to pakkestørrelser. **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.

#### OPGAVE 9 (PRIS OG REKLAME)

- a). **Gennemfør** en passende valgt analyse som undersøger, hvorvidt datamaterialet understøtter en påstand om, at der er forskel på markedsføringsstrategierne for Royal Export og Heineken Premium.

*Vink: Man kan f.eks. basere sin analyse på de to variable *ShPromoRoyal* og *ShPromoHeineken* ELLER man kan i stedet basere sin analyse på de to variable *UgeRoyal* og *UgeHeineken*.*





## DATAKILDE

Datamaterialet i ovenstående opgaver er venligst stillet til rådighed af bryggeriet **Royal Unibrew<sup>TM</sup>** (royalunibrew.com). Virksomheden giver følgende beskrivelse af sig selv:

*We produce, market, sell and distribute quality beverages with focus on branded products within beer, malt beverages and soft drinks as well as cider and long drinks.*

*Our main markets are Denmark, Finland, Italy and Germany as well as Latvia, Lithuania and Estonia. To these should be added the international markets comprising a number of established markets in the Americas region and major cities in Europe and North America as well as emerging markets in for example Africa.*

*In our multi-beverage markets, we offer our customers strong and locally based brands. Based on continuous development and innovation, it is our objective to meet customer demand for quality beverages.*

*In addition to our own brands, we offer licence-based international brands of the PepsiCo and Heineken Groups.*

Datamaterialet er en del af Nielsen ScanTrack Dagligvareindeks<sup>TM</sup> (nielsen.com).

**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

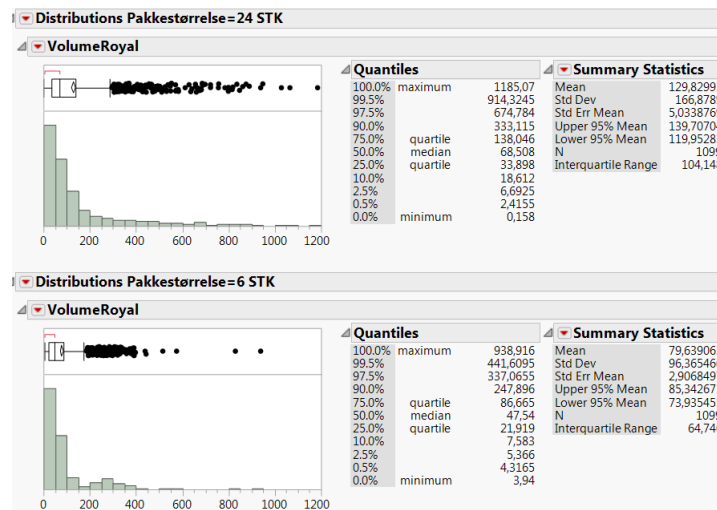
Bemærk at nedenstående løsning...

- 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.



## OPGAVE 1

**1A):** Histogrammerne over fordelingen af salgsvolumen for de to pakkestørrelser viser, at der for begge pakkestørrelser er tale om en meget højreskæv fordeling, hvor der visse uger sælges langt mere end i alle øvrige uger på året. Sådanne uger må umiddelbart betragtes som outliers i forhold til fordelingen af salget i en gennemsnitlig uge på året.



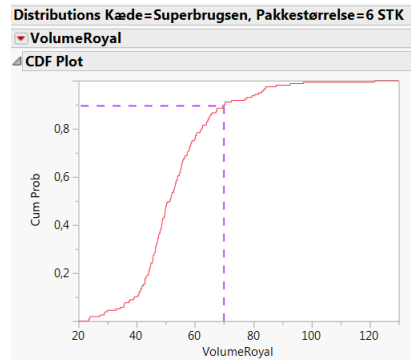
Højreskævheden af fordelingerne fremgår også af de to boxplots, som blandt andet viser, at middelværdierne (24 stk: 12.983 liter; 6 stk: 7.964 liter) er væsentligt større end de tilsvarende medianer (24 stk: 6.851 liter; 6 stk: 4.754 liter), hvilket alt andet lige er en indikation på højreskævhed.

At der er tale om adskillige uger på året med betragteligt større salg end i de øvrige uger ses også af, at standardafvigelserne, som er følsomme overfor outliers, er påvirket af disse uger med meget stort salg, idet standardafvigelserne er væsentligt større (24 stk: 16.688 liter; 6 stk: 9.637 liter) end de tilsvarende interkvartilbredder (24 stk:  $13.805 - 3.390 = 10.415$  liter; 6 stk:  $8.667 - 2.192 = 6.475$  liter), der ikke er følsomme overfor outliers.

**1B):** Median salget af 6 stk Royal Export er på 4.754 liter, svarende til at der i 50% af ugerne på året sælges 4.754 liter eller mindre. På grund af den konstaterede højreskævhed af fordelingen (jf. besvarelsen af delspg. a) ovenfor) er medianen et fint og retvisende mål for centrum af fordelingen, mens middelværdien (7.964 liter) er påvirket af fordelings højreskævhed (idet den er næsten 70% større end medianen). Medianen er derfor umiddelbart det foretrukne mål for centrum i fordelingen.



**1c):** Den kumulerede fordelingsfunktion for salget af Royal Export i 6 stk. pakker i SuperBrugsen viser, at 7.000 liter ca. svarer til 90%-fraktilen i fordelingen. M.a.o. er det kun ca. 10% af ugerne om året (og ikke 25% = hver fjerde uge), at der sælges 7.000 liter eller mere.



**1d):** Et 95%-konfidensinterval for den forventede omsatte mængde af Royal Export i 6 stk. pakker i SuperBrugsen er

[5.123 liter; 5.574 liter]

Confidence Intervals					
Parameter	Estimate	Lower CI	Upper CI	1-Alpha	
Mean	53,48182	51,22935	55,73429	0,950	
Std Dev	14,28822	12,86338	16,07085	0,950	

Med 95% sandsynlighed vil den forventede omsatte mængde således ligge mellem 5.123 liter og 5.574 liter.

Hvis standardafvigelsen af variabelen *VolumeRoyal* stiger, betyder det, at variationen i salget fra uge til uge stiger (dvs. at der alt andet lige bliver større forskelle på salget fra uge til uge). Dermed stiger også usikkerheden omkring forventningen til salget i en given uge, og det betyder, at konfidensintervallet bliver bredere.

**1e):** Test af nulhypotese om at den forventede ugentlige omsatte mængde er 5.250 liter:

**Antagelser:** Observationerne af de enkelte ugers omsætning er indbyrdes uafhængige og approksimativt normalfordelte.

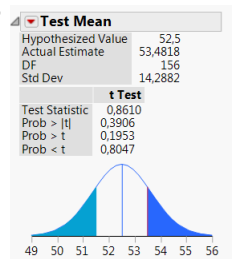
**Hypoteser:**  $H_0 : \mu = 5.250$

$H_a : \mu \neq 5.250$

**Signifikansssh.:**  $p = 0,3906$

**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 ugentlige omsatte mængde Royal Export i 6 stk. pakker i SuperBrugsen er 5.250 liter (svarende til 2.625 stk. six packs).



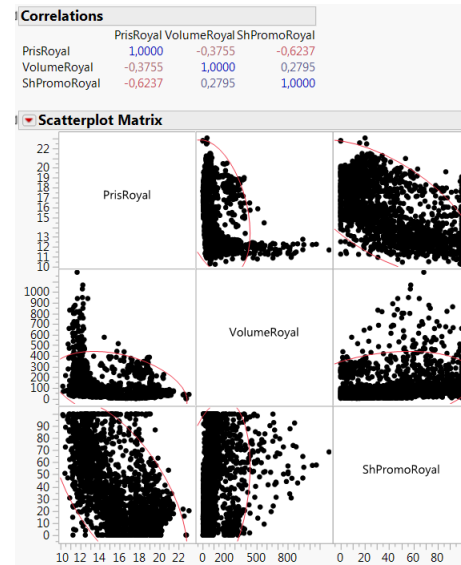


## OPGAVE 2

**2A):** Scatterplottene for Royal Export af henholdsvis pris, volumen og andel solgt til rabatpris kan umiddelbart være en anelse svære at fortolke, fordi de rummer en masse observationer, og det derfor er svært at se hvor i figurerne, de fleste observationer er placeret.

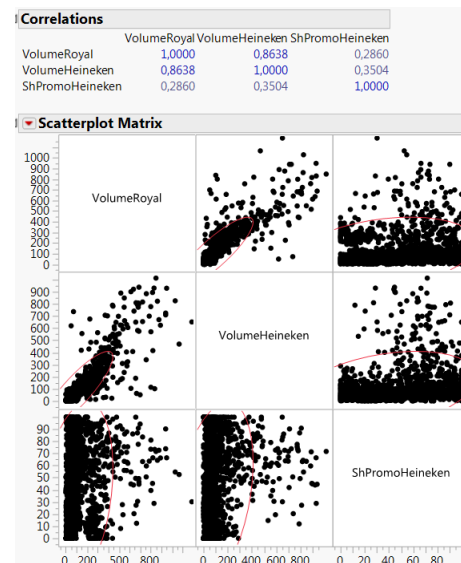
Korrelationen mellem pris og volumen er  $-0,38$ , hvilket indikerer en negativ sammenhæng, således at desto lavere prisen er, desto større mængde sælges der. Korrelationen mellem pris og andel solgt til rabatpris er  $-0,62$ , hvilket indikerer en negativ sammenhæng, således at jo større en del af det samlede salg, der er foregået til rabatpris, desto lavere er varens gennemsnitlige pris (målt som gennemsnit over alle salg i den givne uge). Korrelationen mellem volumen og andel solgt til rabatpris er  $0,28$ , hvilket indikerer en positiv sammenhæng, således at jo større en del der sælges til rabatpris, desto mere sælges der.

For ingen af disse tre korrelationer er fortegnet på korrelationen overraskende i forhold til, hvad man intuitivt skulle forvente.



**2B):** De to scatterplots af andel af Heineken solgt til rabatpris optegnet mod salgsvolumen for henholdsvis Royal Export og Heineken kan umiddelbart være en anelse svære at fortolke, fordi de rummer en masse observationer, og det derfor er svært at se hvor i figurerne, de fleste observationer er placeret. Omvendt viser scatterplottet af salgsvolumen for Royal Export og Heineken en klar positiv lineær sammenhæng (dvs. positiv korrelation).

Korrelationen mellem andelen af Heineken solgt til rabatpris og salgsvolumen for Royal Export er  $0,2860$ , hvilket indikerer en positiv sammenhæng, således at de uger, hvor der sælges meget Heineken til rabatpris, er de samme uger, hvor der også sælges meget Royal Export. Korrelationen mellem andel af Heineken solgt til rabatpris og salgsvolumen for Heineken er  $0,35$ , hvilket indikerer en positiv sammenhæng, således at jo større en del der sælges til rabatpris, desto mere sælges der. Korrelationen mellem salgsvolumen for Heineken og Royal Export er  $0,86$ , hvilket indikerer en klar positiv sammenhæng (som illustreret i scatterplottet), således at de uger, hvor der sælges meget Heineken er de samme uger, hvor der også sælges meget Royal Export.





Korrelation mellem Heinekens salgsvolumen og Heinekens andel solgt til rabatpris er ganske intuitiv (jf. bemærkningerne under delspg. 2a) ovenfor), mens de to øvrige korrelationer umiddelbart kan synes overraskende: Hvorfor sælges der mere Royal Export i samme uger, som der sælges mere Heineken? Og hvorfor sælges der mere Royal Export i de uger, hvor Heineken sælges til rabatpris?

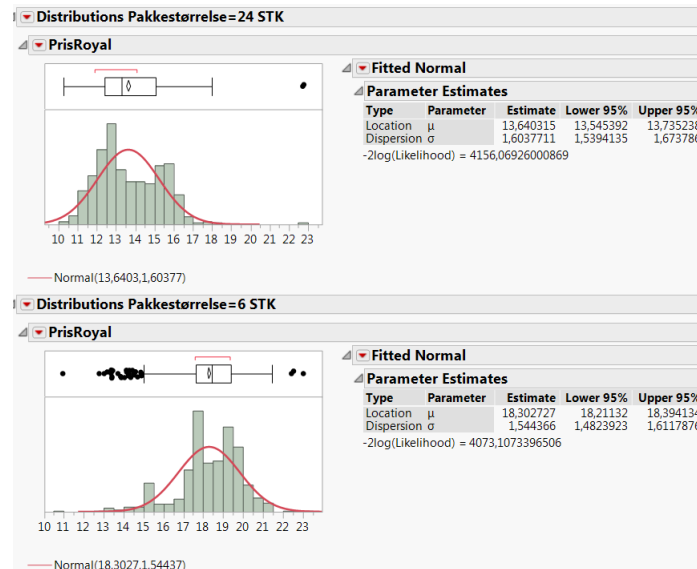
En mulig forklaring på begge korrelationer er, at der i bestemte uger på året sælges væsentligt mere end i alle øvrige uger (f.eks. fordi det om sommeren har været ekstraordinært varmt, eller at salget sker i forbindelse med højtider som påske, pinse, jul og nytår), og at der i disse uger også er en større tendens til at de forskellige øl sælges med rabat (i hvert fald omkring højtiderne). At det er i de samme uger, at begge ølmærker har ekstraordinært stort salg, kan være forklaringen på den positive korrelation.

**2c):** Korrelation mellem to variable udtrykker ikke nødvendigvis en årsagssammenhæng (kausal sammenhæng) mellem de to variable. Af de tre korrelationer i delspg 2b) er det kun korrelationen mellem andelen af Heineken solgt til rabatpris og salgsvolumen for Heineken, der har en naturlig kausal fortolkning. For de øvrige to korrelationers vedkommende gælder, at den positive korrelation formentlig er drevet af en underliggende faktor i form af den tidsmæssige variation henover år, som beskrevet i besvarelsen af delspg. 2b) ovenfor.

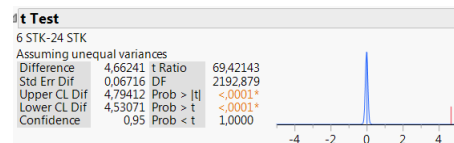


## OPGAVE 3

**3A):** Histogrammerne over fordelingerne af gennemsnitsprisen på Royal Export for de to pakkestørrelser viser, at begge fordelinger er nogenlunde klokkeformede, hvilket er i god overensstemmelse med normalfordelingsantagelsen. Histogrammerne viser også, at der i begge fordelinger er enkelte outliers, hvilket ikke stemmer helt overens med normalfordelingsantagelsen. Eftersom der kun er tale om ganske få outliers, er det imidlertid ikke noget, som taler væsentligt imod, at normalfordelingen giver en fornuftig beskrivelse af prisen på hver af de to pakkestørrelser.



**3B):** Test af nulhypotese om at den forventede pris pr. liter Royal Export er den samme for 6 stk. og 24 stk. pakker:



Antagelser:

Observationerne af de enkelte ugers gennemsnitlige salgpris er indbyrdes uafhængige og approksimativt normalfordelte; normalfordelte indenfor hver pakkestørrelse.

Hypoteser:

$$H_0 : \mu_{6 \text{ stk.}} = \mu_{24 \text{ stk.}}$$

$$H_a : \mu_{6 \text{ stk.}} \neq \mu_{24 \text{ stk.}}$$

Signifikanssssh.:

$$p < 0,0001$$

Tekn. konklusion:

Signifikanssandsynligheden er mindre end 0,05 og dermed forkastes  $H_0$ .

Let konklusion:

Der er ikke belæg for en påstand om, at den forventede pris pr. liter Royal Export er den samme for 6 stk. og 24 stk. pakker. På baggrund af histogrammerne i delspg. a) er den konklusion ikke overraskende, eftersom den forventede pris pr. liter (centrum af histogrammet) er ca. 13-14 kr. (helt præcist: 13,64 kr.) for 24 stk. pakker og 18-19 kr. (helt præcist: 18,30 kr.) for 6 stk. pakker. Så desto større pakke man køber, desto billigere er det, hvilket næppe er en særligt overraskende konklusion.



**3c):** Et 95%-konfidensinterval for forskellen i den forventede pris pr. liter Royal Export for hhv. 6 stk. og 24 stk. pakker er (jf. JMP-outputtet på forrige side)

[4,53 kr.; 4,79 kr.]

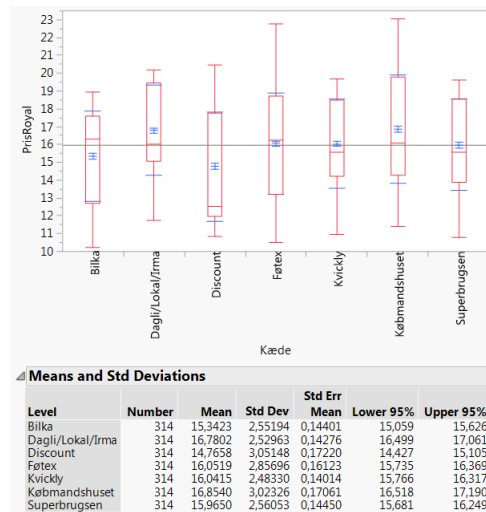
Med 95% sandsynlighed vil den forventede forskel i prisen pr. liter Royal Export for hhv. 6 stk. og 24 stk. pakker således ligge mellem 4,53 kr. og 4,79 kr. Løst sagt: En six pack Royal Export vil forventeligt være mellem 4,53 kr. og 4,79 kr. dyrere pr. liter end en ramme Royal Export.





#### OPGAVE 4

**4A):** Boxplot samt middelværdi og standardafvigelse for fordelingen af den gennemsnitlige pris pr. liter Royal Export for hver af de betragtede supermarkeds-kæder:





**4b):** Test af nulhypotese om at den forventede pris pr. liter Royal Export er den samme i de 7 forskellige supermarkeds kæder:

Oneway Anova					
Summary of Fit					
Rsquare		0,059482			
Adj Rsquare		0,056906			
Root Mean Square Error		2,732043			
Mean of Response		15,97152			
Observations (or Sum Wgts)		2198			
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Kæde	6	1034,273	172,379	23,0945	<,0001*
Error	2191	16353,753	7,464		
C. Total	2197	17388,026			
Means for Oneway Anova					
Level	Number	Mean	Std Error	Lower 95%	Upper 95%
Bilka	314	15,3423	0,15418	15,040	15,645
Dagli/Lokal/Irma	314	16,7802	0,15418	16,478	17,083
Discount	314	14,7658	0,15418	14,463	15,068
Føtex	314	16,0519	0,15418	15,750	16,354
Kvickly	314	16,0415	0,15418	15,739	16,344
Købmandshuset	314	16,8540	0,15418	16,552	17,156
Superbrugsen	314	15,9650	0,15418	15,663	16,267

Std Error uses a pooled estimate of error variance

**Antagelser:** Observationerne af de ugentlige gennemsnitspriser i de enkelte supermarkeds kæder er indbyrdes uafhængige og approksimativt normalfordelte; normalfordelte indenfor hver supermarkeds kæde og med samme std.afvigelse for alle supermarkeds kæder. Hvorvidt antagelsen om normalfordeling indenfor hver supermarkeds kæde er opfyldt kan ikke vurderes ud fra de i delspg. a) optegnede box plots, men antagelsen om samme std.afvigelse ser iflg. boxplottene ud til at være nogenlunde opfyldt.

**Hypoteser:**  $H_0 : \mu_{\text{Bilka}} = \mu_{\text{Dagli/Lokal/Irma}} = \mu_{\text{Discount}} = \mu_{\text{Føtex}} = \mu_{\text{Kvickly}} = \mu_{\text{Købmandshuset}} = \mu_{\text{SuperBrugsen}}$   
 $H_a : \text{Mindst to af de forventede værdier } \mu_{\text{Bilka}}, \mu_{\text{Dagli/Lokal/Irma}}, \mu_{\text{Discount}}, \mu_{\text{Føtex}}, \mu_{\text{Kvickly}}, \mu_{\text{Købmandshuset}}, \mu_{\text{SuperBrugsen}} \text{ er ikke ens}$

**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 den forventede pris pr. liter Royal Export er den samme i alle 7 supermarkeds kæder. Konklusionen er ikke overraskende på baggrund af resultaterne i delspg. a), som viser at den gennemsnitlige pris i discountbutikker (14,77 kr.) er ca. 2 kr. lavere end i eksempelvis Dagli/Lokal/Irma (16,78 kr.) og i Købmandshuset (16,85 kr.).



**4c):** Af resultaterne i delspg. a) fremgår det, at de tre supermarkeds kæder, hvis gennemsnitlige pris pr. liter Royal Export ligger tættest på hinanden, er Føtex (16,05 kr.), Kvickly (16,04 kr.) og SuperBrugsen (15,97 kr.). Det er derfor naturligt at undersøge en nulhypotese om prisen i disse tre supermarkeds kæder.

Test af nulhypotese om at den forventede pris pr. liter Royal Export er den samme i Føtex, Kvickly og SuperBrugsen:

Oneway Anova					
Summary of Fit					
Rsquare			0,000216		
Adj Rsquare			-0,00191		
Root Mean Square Error			2,63852		
Mean of Response			16,01945		
Observations (or Sum Wgts)			942		
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Kæde	2	1,4148	0,70741	0,1016	0,9034
Error	939	6537,1202	6,96179		
C. Total	941	6538,5351			
Means for Oneway Anova					
Level	Number	Mean	Std Error	Lower 95%	Upper 95%
Føtex	314	16,0519	0,14890	15,760	16,344
Kvickly	314	16,0415	0,14890	15,749	16,334
Superbrugsen	314	15,9650	0,14890	15,673	16,257
Std Error uses a pooled estimate of error variance					

- Antagelser:** Observationerne af de ugentlige gennemsnitspriser i de 3 supermarkeds kæder er indbyrdes uafhængige og approksimativt normalfordelte; normalfordelte indenfor hver supermarkeds kæde og med samme std.afvigelse for alle 3 supermarkeds kæder.
- Hypoteser:**  $H_0 : \mu_{Føtex} = \mu_{Kvickly} = \mu_{SuperBrugsen}$   
 $H_a : \text{Mindst to af de forventede værdier } \mu_{Føtex}, \mu_{Kvickly}, \mu_{SuperBrugsen} \text{ er ikke ens}$
- Signifikansssh.:**  $p = 0,9034$
- 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 pris pr. liter Royal Export er den samme i alle tre supermarkeds kæder. Løst sagt: Der er ikke forskel på prisen på Royal Export i de tre supermarkeds kæder.



## OPGAVE 5

**5A):** Variablene i den reducerede regressionsmodel (ved et signifikansniveau på 5%) er *PrisRoyal*, *Kæde*, *Pakkestørrelse*, *VolumeCarlsberg*, *VolumeHeineken*, *ShPromoCarlsberg*, *ShPromoHeineken*, *ShPromoRoyal*.

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
PrisRoyal	1	1	123128.5	39.0621	<.0001*
Kæde	6	6	890919.2	47.1068	<.0001*
Pakkestørrelse	1	1	45867.3	14.5512	0.0001*
VolumeCarlsberg	1	1	204677.8	64.9333	<.0001*
VolumeHeineken	1	1	6160157.5	1954.289	<.0001*
ShPromoCarlsberg	1	1	66823.4	21.1995	<.0001*
ShPromoHeineken	1	1	12420.9	3.9405	0.0473*
ShPromoRoyal	1	1	59779.2	18.9648	<.0001*

**5B):** Eftersom vi fjerner variable fra regressionsmodellen, så længe deres signifikanssandsynlighed er større end signifikansniveauet (der normalt sættes til 5%), så betyder det, at jo lavere signifikanssandsynlighed vi anvender, desto flere variable vil vi alt andet lige kunne fjerne fra modellen. Såfremt vi anvender et signifikansniveau på 1% frem for 5%, vil vi derfor få samme eller færre variable i den reducerede model.

**5C):**

Parameterestimaterne hørende til variablen *Kæde* er...

- $\hat{\beta}_{Kæde,Bilka} = -36,6014$
- $\hat{\beta}_{Kæde,Dagli/Lokal/Irma} = 7,2128$
- $\hat{\beta}_{Kæde,Discount} = 89,8785$
- $\hat{\beta}_{Kæde,Føtex} = -11,2475$
- $\hat{\beta}_{Kæde,Kvickly} = -24,9031$
- $\hat{\beta}_{Kæde,Købmandshuset} = -9,6222$
- $\hat{\beta}_{Kæde,SuperBrugsen} = 0$

Resultatet af den estimerede model:

Indicator Function Parameterization						
Term	Estimate	Std Error	t Ratio	Prob> t	Lower 95%	Upper 95%
Intercept	154.17538	20.66591	7.46	<.0001*	113.64848	194.70229
PrisRoyal	-6.630722	1.060921	-6.25	<.0001*	-8.711241	-4.550202
Kæde[Bilka]	-36.60137	4.768871	-7.68	<.0001*	-45.95337	-27.24937
Kæde[Dagli/Lokal/Irma]	7.212763	4.562579	1.58	0.1141	-1.734686	16.160212
Kæde[Discount]	89.878451	7.137128	12.59	<.0001*	75.882181	103.87472
Kæde[Føtex]	-11.24751	4.595425	-2.45	0.0145*	-20.25937	-2.23565
Kæde[Kvickly]	-24.90309	4.556246	-5.47	<.0001*	-33.83812	-15.96806
Kæde[Købmandshuset]	-9.622161	4.58033	-2.10	0.0358*	-18.60442	-0.639901
Pakkestørrelse[24 STK]	-18.99849	4.980456	-3.81	0.0001*	-28.76542	-9.231562
VolumeCarlsberg	0.1206722	0.014975	8.06	<.0001*	0.091305	0.1500394
VolumeHeineken	0.6174935	0.013968	44.21	<.0001*	0.5901013	0.6448857
ShPromoCarlsberg	-0.163418	0.035493	-4.60	<.0001*	-0.233021	-0.093815
ShPromoHeineken	-0.125027	0.062984	-1.99	0.0473*	-0.248541	-0.001512
ShPromoRoyal	0.2946507	0.06766	4.35	<.0001*	0.1619654	0.4273359

Det betyder, at det forventede ugentlige salg af Royal Export af en given pakkestørrelse (6 eller 24 stk.) alt andet lige er...

- 3.660 liter lavere i Bilka end i SuperBrugsen
- 721 liter større i Dagli/Lokal/Irma end i SuperBrugsen
- 8.988 liter større i discountbutikkerne end i SuperBrugsen
- 1.125 liter lavere i Føtex end i SuperBrugsen
- 2.490 liter lavere i Kvickly end i SuperBrugsen
- 962 liter lavere i Købmandshuset end i SuperBrugsen



Parameterestimatet hørende til variablen...

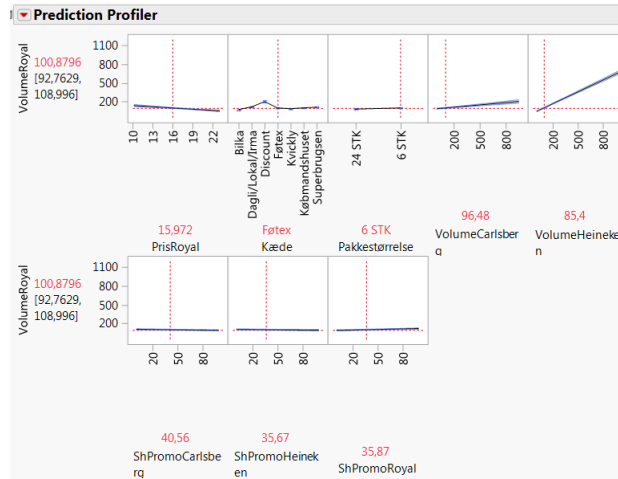
- **PrisRoyal** er  $\hat{\beta}_{\text{PrisRoyal}} = -6,6307$ , således at det forventede ugentlige salg af Royal Export af en given pakkestørrelse (6 eller 24 stk.) falder med 663 liter, hvis den gennemsnitlige pris stiger med 1 kr. pr. liter.
- **Pakkestørrelsen** er  $\hat{\beta}_{\text{Pakkestørrelse}, 24 \text{ stk}} = -18,9985$ , hvis vi ser på salget af 24 stk. pakker, og  $\hat{\beta}_{\text{Pakkestørrelse}, 6 \text{ stk}} = 0$ , hvis vi ser på salget af 6 stk. pakker. Det forventede ugentlige salg af Royal Export i 24 stk pakker i en given supermarkeds-kæde er således alt andet lige 1.900 liter lavere end det tilsvarende salg af 6 stk. pakker.
- **VolumeCarlsberg** er  $\hat{\beta}_{\text{VolumeCarlsberg}} = 0,1207$ , således at det forventede ugentlige salg af Royal Export af en given pakkestørrelse i en given supermarkeds-kæde stiger med 12 liter, hvis salget af Carlsberg Pilsner i samme uge stiger med 100 liter.
- **VolumeHeineken** er  $\hat{\beta}_{\text{VolumeHeineken}} = 0,6175$ , således at det forventede ugentlige salg af Royal Export af en given pakkestørrelse i en given supermarkeds-kæde stiger med 62 liter, hvis salget af Heineken Premium i samme uge stiger med 100 liter.
- **ShPromoCarlsberg** er  $\hat{\beta}_{\text{ShPromoCarlsberg}} = -0,1634$ , således at det forventede ugentlige salg af Royal Export af en given pakkestørrelse i en given supermarkeds-kæde falder med 16 liter, hvis andelen af Carlsberg Pilsner solgt til rabatpris stiger med 1 procent.
- **ShPromoHeineken** er  $\hat{\beta}_{\text{ShPromoHeineken}} = -0,1250$ , således at det forventede ugentlige salg af Royal Export af en given pakkestørrelse i en given supermarkeds-kæde falder med 13 liter, hvis andelen af Heineken Premium solgt til rabatpris stiger med 1 procent.
- **ShPromoRoyal** er  $\hat{\beta}_{\text{ShPromoRoyal}} = 0,2947$ , således at det forventede ugentlige salg af Royal Export af en given pakkestørrelse i en given supermarkeds-kæde stiger med 29 liter, hvis andelen af Royal Export solgt til rabatpris stiger med 1 procent.

**5d):** Et 95%-konfidensinterval for betydningen af en prisstigning på 1 kr. pr. liter Royal Export er (jf. JMP-outputtet på forrige side)  $[-8,7112; -4,5502]$ . Med 95% sandsynlighed vil det forventede ugentlige salg af Royal Export af en given pakkestørrelse i en given supermarkeds-kæde dermed falde med mellem 455 og 871 liter, såfremt prisen pr. liter Royal Export stiger med 1 kr.

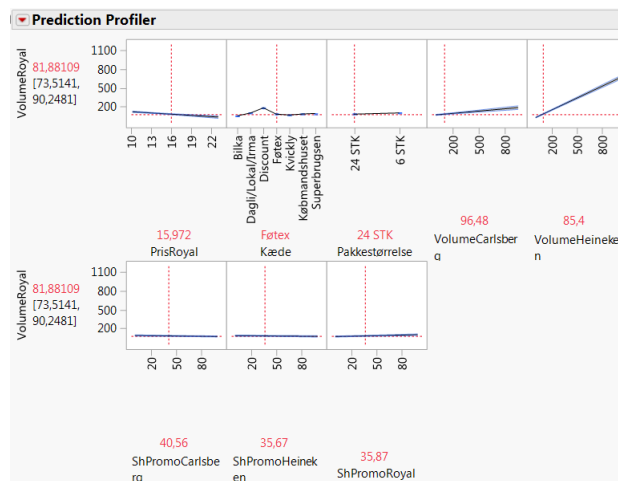
**5e):** På baggrund af tallene fra besvarelsen af delspg. d) ovenfor ses det, at det forventede ugentlige salg af Royal Export af en given pakkestørrelse i en given supermarkeds-kæde vil falde med mellem  $4,5502 \cdot 0,5 = 2,2751$  og  $8,7112 \cdot 0,5 = 4,3556$  (målt i 100 liter), såfremt prisen pr. liter Royal Export stiger med 0,50 kr. Det efterspurgte 95%-konfidensinterval bliver dermed  $[-4,3556; -2,2751]$ .



5f): Det forventede ugentlige salg af Royal Export i 6 stk. pakker i Føtex (når alle øvrige forklarende variable sættes lig deres respektive middelværdier) er 10.088 liter.

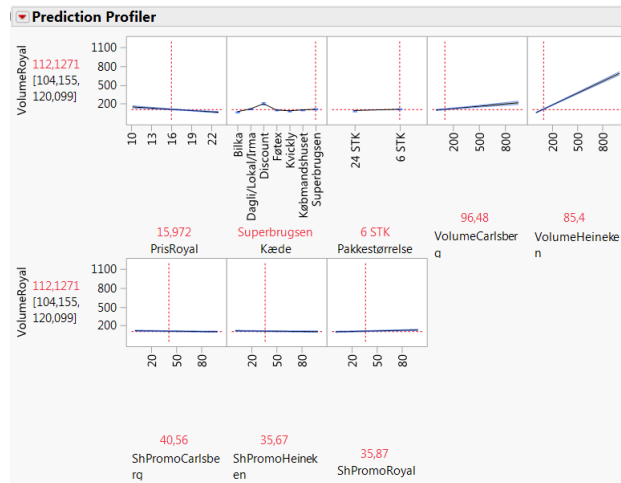


5g): Det forventede ugentlige salg af Royal Export i 24 stk. pakker i Føtex (når alle øvrige forklarende variable sættes lig deres respektive middelværdier) er 8.188 liter. Det forventede salg falder således med 10.088 - 7.351 = 1.900 liter (idet  $\hat{\beta}_{\text{Pakkestørrelse}, 24 \text{ stk}} = -18,9985$ ).





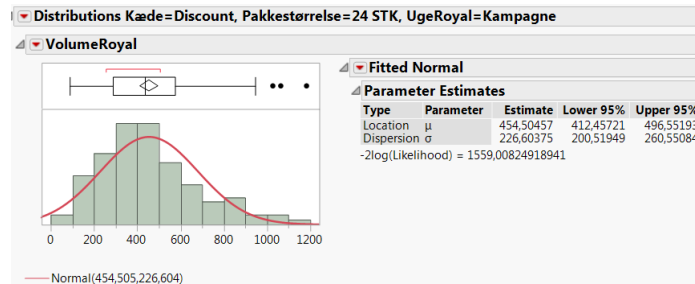
Det forventede ugentlige salg af Royal Export i 6 stk. pakker i SuperBrugsen (når alle øvrige forklarende variable sættes lig deres respektive middelværdier) er 11.213 liter. Det forventede salg stiger således med  $11.213 - 10.088 = 1.125$  liter (idet  $\hat{\beta}_{K\ddot{a}de, F\ddot{o}tex} = -11,2475$  og  $\hat{\beta}_{K\ddot{a}de, SuperBrugsen} = 0$ ).





## OPGAVE 6

**6A):** Histogram samt middelværdi og standardafvigelse for fordelingen af salget af Royal Export i 24 stk. pakker i discountbutikker i en kampagneuge:



**6B):** Sandsynligheden for at der sælges mindst 20.000 liter Royal Export i 24 stk. pakker i discountbutikker i en kampagneuge er

$$P(\text{VolumeRoyal} > 200) = 1 - P(\text{VolumeRoyal} \leq 200) = 1 - 13,1\% = 86,7\%$$

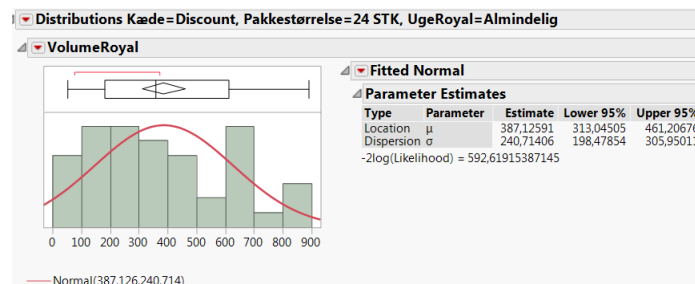
hvor sandsynligheden er beregnet i en normalfordeling med forventet værdi og std.afvigelse som angivet i delspg. a).

Sandsynligheden for at der sælges mellem 20.000 og 70.000 liter Royal Export i 24 stk. pakker i discountbutikker i en kampagneuge er

$$P(200 \leq \text{VolumeRoyal} \leq 700) = P(\text{VolumeRoyal} \leq 700) - P(\text{VolumeRoyal} \leq 200) = 86,1\% - 13,1\% = 73,0\%$$

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

**6C):** Middelværdi og standardafvigelse for fordelingen af salget af Royal Export i 24 stk. pakker i discountbutikker i en almindelig uge:



Sandsynligheden for at der sælges mellem 20.000 og 70.000 liter Royal Export i 24 stk. pakker i discountbutikker i en almindelig uge er

$$P(200 \leq \text{VolumeRoyal} \leq 700) = P(\text{VolumeRoyal} \leq 700) - P(\text{VolumeRoyal} \leq 200) = 90,3\% - 21,9\% = 68,5\%$$

hvor sandsynligheden er beregnet i en normalfordeling med forventet værdi og std.afvigelse som angivet i JMP-outputtet ovenfor.





## OPGAVE 7

7A): Krydstabel mellem supermarkeds-kæde og hvorvidt der er tale om en kampagne eller ej:

	UgeRoyal		Total
	Almindelig	Kampagne	
Count			
Total %			
Col %			
Row %			
Bilka	214	100	314
	9,74	4,55	14,29
	14,12	14,66	
	68,15	31,85	
Dagli/Lokal/Irma	243	71	314
	11,06	3,23	14,29
	16,03	10,41	
	77,39	22,61	
Discount	184	130	314
	8,37	5,91	14,29
	12,14	19,06	
	58,60	41,40	
Føtex	216	98	314
	9,83	4,46	14,29
	14,25	14,37	
	68,79	31,21	
Kvickly	220	94	314
	10,01	4,28	14,29
	14,51	13,78	
	70,06	29,94	
Købmandshuset	229	85	314
	10,42	3,87	14,29
	15,11	12,46	
	72,93	27,07	
Superbrugsen	210	104	314
	9,55	4,73	14,29
	13,85	15,25	
	66,88	33,12	
Total	1516	682	2198
	68,97	31,03	

Andelen af kampagneuger i hver af de 7 supermarkeds-kæder er:

$$\hat{p}_{\text{Bilka}} = 31,85\% \quad \hat{p}_{\text{Dagli/Lokal/Irma}} = 22,61\% \quad \hat{p}_{\text{Discount}} = 41,40\% \quad \hat{p}_{\text{Føtex}} = 31,21\%$$

$$\hat{p}_{\text{Kvickly}} = 29,94\% \quad \hat{p}_{\text{Købmandshuset}} = 27,07\% \quad \hat{p}_{\text{SuperBrugsen}} = 33,12\%$$

Dagli/Lokal/Irma er med en andel på 22,61% den af de 7 supermarkeds-kæder med lavest andel kampagneuger. Løst sagt: Dagli/Lokal/Irma er den kæde, der sjældnest har tilbud på Royal Export, nemlig gennemsnitligt lidt mindre end hver fjerde uge om året.



**7b):** Test af nulhypotese om at der er uafhængighed mellem supermarkeds-kæde og hvorvidt der er kampagne for Royal Export eller ej:

- Antagelser:** Observationerne af de ugentlige omsætninger er indbyrdes uafhængige. Endvidere er det forventede antal observationer i hver af krydstabellens celler større end 5. Sidstnævnte antagelse er klart opfyldt, da alle celler har et forventet antal observationer på 97 eller mere.
- Hypoteser:**  $H_0$  : Uafhængighed mellem **Kæde** og **UgeRoyal**  
 $H_a$  : Afhængighed mellem **Kæde** og **UgeRoyal**
- Signifikanssssh.:**  $p < 0,0001$
- Tekn. konklusion:** Signifikanssandsynligheden er mindre end 0,05 og dermed forkastes  $H_0$ .
- Let konklusion:** Der er ikke belæg for en påstand om, at der er uafhængighed mellem hvilken supermarkeds-kæde omsætningen stammer fra, og hvorvidt der er tale om en kampagneuge eller ej. M.a.o. er der belæg for en påstand om, at der er en sammenhæng (afhængighed) mellem sandsynligheden for at omsætningen stammer fra en kampagneuge, og hvilken kæde omsætningen stammer fra. Løst sagt: Der er forskel på hvor ofte de forskellige supermarkeds-kæder har kampagne for Royal Export.

Af krydstabellen i delspg. a) fremgår det, at hyppigheden af kampagner for Royal Export varierer fra knap hver fjerde uge (i Dagli/Lokal/Irma) til hver anden-tredje uge (i discountbutikker), og på den baggrund er konklusionen ikke videre overraskende.

**Contingency Analysis of UgeRoyal By Kæde**

**Contingency Table**

	UgeRoyal		
	Almindelig	Kampagne	Total
Count			
Total % Expected			
Bilka	214	100	314
	9,74	4,55	14,29
	216,571	97,4286	
Dagli/Lokal/Irma	243	71	314
	11,06	3,23	14,29
	216,571	97,4286	
Discount	184	130	314
	8,37	5,91	14,29
	216,571	97,4286	
Fotex	216	98	314
	9,83	4,46	14,29
	216,571	97,4286	
Kvickly	220	94	314
	10,01	4,28	14,29
	216,571	97,4286	
Købmandshuset	229	85	314
	10,42	3,87	14,29
	216,571	97,4286	
Superbrugsen	210	104	314
	9,55	4,73	14,29
	216,571	97,4286	
Total	1516	682	2198
	68,97	31,03	

**Tests**

	N	DF	-LogLik	RSquare (U)
	2198	6	14,652538	0,0108

Test	ChiSquare	Prob> ChiSq
Likelihood Ratio	29,305	<.0001*
Pearson	29,401	<.0001*



**7c):** Test af nulhypotese om at der er uafhængighed mellem supermarkedskæde og hvorvidt der er kampagne for Royal Export eller ej (når vi ser bort fra kæderne Dagli/Lokal/Irma og Discount):

- Antagelser:** Observationerne af de ugentlige omsætninger er indbyrdes uafhængige. Endvidere er det forventede antal observationer i hver af krydstabellens celler større end 5. Sidstnævnte antagelse er klart opfyldt, da alle celler har et forventet antal observationer på 96 eller mere.
- Hypoteser:**  $H_0$  : Uafhængighed mellem **Kæde** og **UgeRoyal**  
 $H_a$  : Afhængighed mellem **Kæde** og **UgeRoyal**
- Signifikansssh.:**  $p = 0,5364$
- Tekn. konklusion:** Signifikanssandsynligheden er på 0,5364 og dermed større end 0,05, og dermed kan  $H_0$  ikke forkastes.
- Let konklusion:** Vi kan ikke afvise hypotesen om, at den forventede andel af kampagneuger er den samme i de 5 betragtede supermarkedskæder. Løst sagt: Supermarkedskæderne Bilka, Føtex, Kvickly, Købmandshuset og SuperBrugsen har lige ofte tilbud på Royal Export.

**Contingency Analysis of UgeRoyal By Kæde**

**Contingency Table**

	UgeRoyal		
	Almindelig	Kampagne	Total
Count			
Total % Expected			
Bilka	214 13,63 217,8	100 6,37 96,2	314 20,00
Føtex	216 13,76 217,8	98 6,24 96,2	314 20,00
Kvickly	220 14,01 217,8	94 5,99 96,2	314 20,00
Købmandshuset	229 14,59 217,8	85 5,41 96,2	314 20,00
Superbrugsen	210 13,38 217,8	104 6,62 96,2	314 20,00
Total	1089 69,36	481 30,64	1570

**Tests**

	N	DF	-LogLike	RSquare (U)
	1570	4	1,5794318	0,0016

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	3,159	0,5316
Pearson	3,129	0,5364



## OPGAVE 8

**8A):** Krydstabel mellem pakkestørrelse og hvorvidt der er tale om en kampagneuge eller ej:

Contingency Analysis of UgeRoyal By Pakkestørrelse				
Contingency Table				
	Count	UgeRoyal		Total
		Almindelig	Kampagne	
Pakkestørrelse	Total %			
	Col %			
24 STK	Row %			
		570	529	1099
6 STK		25,93	24,07	50,00
		37,60	77,57	
		51,87	48,13	
		946	153	1099
		43,04	6,96	50,00
		62,40	22,43	
Total		86,08	13,92	
		1516	682	2198
		68,97	31,03	

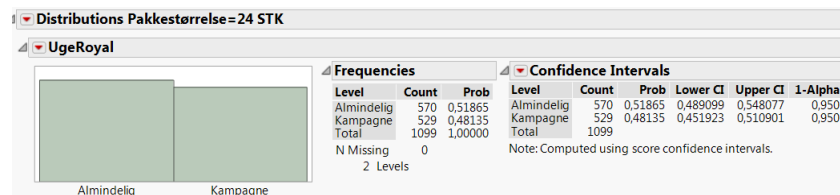
Andelen af kampagneuger for hver af de to pakkestørrelser er:

$$\hat{p}_{6stk} = 13,92\% \quad \hat{p}_{24stk} = 48,13\%$$

**8B):** Et 95%-konfidensinterval for andelen af kampagneuger på et år for Royal Export i 24 stk. pakker er

$$[0,4519; 0,5109]$$

Med 95% sandsynlighed vil den forventede årlige andel kampagneuger for Royal Export i 24 stk. pakker ligge mellem 45,2% og 51,1%. Løst sagt: En ramme Royal Export vil forventeligt være på tilbud ca. hver anden uge om året (i en eller flere supermarkeds kæder).



Antagelserne bag beregningen af konfidensintervallet er, at de ugentlige observationer af om der i en given uge er kampagne for Royal Export eller ej skal være indbyrdes uafhængige. Desuden skal datamaterialet indeholde mindst 15 observationer af uger uden kampagne og mindst 15 af uger med kampagne.



**8c):** Test af nulhypotese om at den forventede andel kampagneuger er den samme for 6 og 24 stk. pakker Royal Export:

Contingency Analysis of UgeRoyal By Pakkestørrelse				
Contingency Table				
	UgeRoyal			
	Almindelig	Kampagne	Total	
Count				
Total %				
Col %				
Row %				
Expected				
Pakkestørrelse 24 STK	570	529	1099	
	25.93	24.07	50.00	
	37.60	77.57		
	51.87	48.13		
	758	341		
6 STK	946	153	1099	
	43.04	6.96	50.00	
	62.40	22.43		
	86.08	13.92		
	758	341		
Total	1516	682	2198	
	68.97	31.03		

Two Sample Test for Proportions			
Description	Proportion Difference	Lower 95%	Upper 95%
P(Almindelig 24 STK)-P(Almindelig 6 STK)	-0.34213	-0.37744	-0.30558
Adjusted Wald Test		Prob	
P(Almindelig 24 STK)-P(Almindelig 6 STK) ≥ 0		1.0000	
P(Almindelig 24 STK)-P(Almindelig 6 STK) ≤ 0		<.0001*	
P(Almindelig 24 STK)-P(Almindelig 6 STK) = 0		<.0001*	

Antagelser:

De ugentlige observationer af om der i en given uge i en given supermarkedskæde for en given pakkestørrelse er kampagne for Royal Export eller ej er indbyrdes uafhængige. Endvidere er det forventede antal observationer i hver af krydstabelens celler større end 15. Sidstnævnte antagelse er opfyldt, da alle celler har et forventet antal observationer på 341 eller mere.

Hypoteser:

$$H_0 : p_{6stk.} = p_{24stk.}$$

$$H_a : p_{6stk.} \neq p_{24stk.}$$

Signifikanssssh.:

$$p < 0,0001$$

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 kampagneuger er den samme for de to pakkestørrelser.

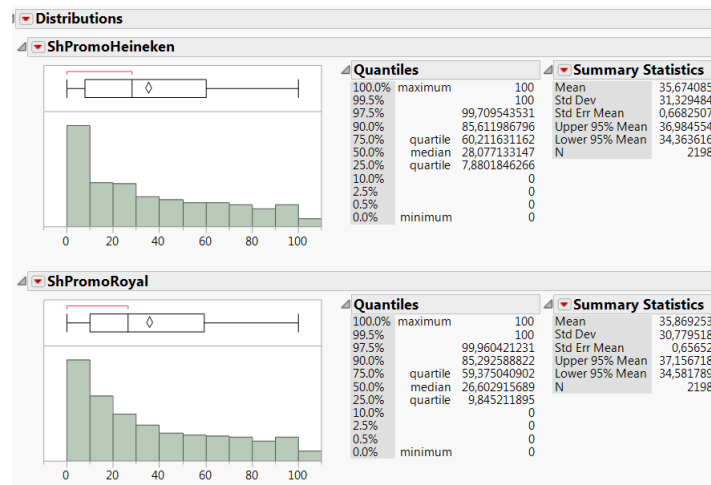
På baggrund af resultaterne i delspg. a) er den konklusion ikke overraskende, eftersom Royal Export i 6 stk. pakker i gennemsnit er på tilbud ca. hver syvende uge (= 1/13,92%), mens Royal Export i 24 stk. pakker i gennemsnit er på tilbud ca. hver anden uge (= 1/48,13%).



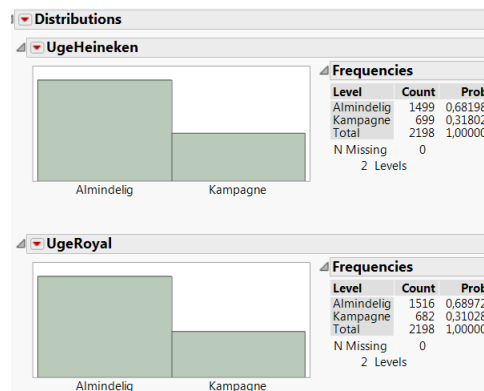
### OPGAVE 9

Spørgsmålet om hvorvidt der er forskel på markedsføringsstrategierne for Royal Export og Heineken Premium kan besvares på flere forskellige måder.

Man kan eksempelvis lave en deskriptiv analyse af de to variable *ShPromoHeineken* og *ShPromoRoyal* og på baggrund af den konkludere, at der ikke umiddelbart ser ud til at være iøjnefaldende forskelle på variablene. Der ser således ikke umiddelbart ud til at være forskel på de to markedsføringsstrategier



Alternativt kan man gøre det samme blot for variablene *UgeHeineken* og *UgeRoyal*, hvilket umiddelbart leder til samme konklusion som ovenfor.





Endelig er en tredje mulighed at lave et test for om to andele er ens, når vi sammenligner variablene *UgeHeineken* og *UgeRoyal*. Her forkastes nulhypotesen, hvilket indikerer at der er en vis sammenhæng mellem de to variable, altså en vis lighed mellem de to markedsføringsstrategier.

Contingency Analysis of UgeHeineken By UgeRoyal				
Contingency Table				
UgeRoyal		UgeHeineken		Total
		Almindelig	Kampagne	
Almindelig	Count	1362	154	1516
	Total %	61.97	7.01	68.97
	Col %	90.86	22.03	
	Row %	89.84	10.16	
	Expected	1033.89	482.113	
Kampagne	Count	137	545	682
	Total %	6.23	24.80	31.03
	Col %	9.14	77.97	
	Row %	20.09	79.91	
	Expected	465.113	216.887	
Total	Count	1499	699	2198
	Total %	68.20	31.80	
Two Sample Test for Proportions				
Description	Proportion Difference	Lower 95%	Upper 95%	
P(Almindelig Almindelig)-P(Almindelig Kampagne)	0.697537	0.662426	0.729849	
Adjusted Wald Test		Prob		
P(Almindelig Almindelig)-P(Almindelig Kampagne) ≥ 0		<.0001*		
P(Almindelig Almindelig)-P(Almindelig Kampagne) ≤ 0		1.0000		
P(Almindelig Almindelig)-P(Almindelig Kampagne) = 0		<.0001*		

PS. Principielt kan man også lave flere andre former for hypotesetest, men det kræver i så fald at man selv udregner teststørrelse og signifikanssandsynlighed, fordi testene ikke kan laves i JMP ud fra den måde data er organiseret på i den givne datafil. Alternativt skal man selv omorganisere data og herefter lave testet. Ingen af disse ting er dog noget, som forventes i forbindelse med besvarelsen af eksamensopgaven.