Tentamen i: **Finansiell statistik**

Examinator: Mikael Möller

Hjälpmedel: Miniräknare (med, i förekommande fall, tomt programmerbart minne).
Utdelad formelsamling samt *Tabeller över statistiska fördelningar*.

Tentamensdag: 030116

Skrivningstid: 5 timmar

Återlämningsdag: Fredag 31:a Januari kl. $13^{00} - 13^{30}$ i B705 plan 7.
Därefter kan skrivningarna hämtas på expeditionen plan 7 i B-huset.

Betygskrav: För godkänt krävs preliminärt minst 30 poäng varav minst 15 poäng på varje delavsnitt.
För väl godkänt krävs preliminärt minst 50 poäng varav minst 20 poäng på varje delavsnitt.

Lösningarna skall vara **väl motiverade** och lätta att följa. Saknas motivering görs avdrag. Om antaganden och analys ej klart framgår kan det bli fullt avdrag.
Om mer än en lösning inlämnas på en uppgift rättas den sämsta lösningen.
Endast ett tal:s lösning per pappersark.
Tabellsamlingen får inte innehålla några anteckningar.

Observera: De studenter **som tidigare ansökt** om på statistiska institutionen **och beviljats** rätt att tillgodoräkna sig del av kursen får välja mellan två alternativ. Valet skall meddelas skrivvakten innan tentamen påbörjas.
Alternativen är att

a. göra hela tentan. Skrivtiden är då 5 timmar. Märk kuvertet "5 timmar".

b. endats göra uppgifterna under rubriken *Tillämpad statistik* (tal 4, 5 och 6). Skrivtiden är då 2.5 timmar. I detta fall är maximal poäng 30 och för godkänt krävs minst 15 poäng. *Betyget väl godkänt delas inte ut i detta fall.* Märk kuvertet ”2.5 timmar”.

Lycka till!

Sannolikhetslära och inferens

Lösning 1: Vi antager att villkoren för en Bernoulliprocess är uppfyllda dvs att den underliggande sannolikhetsstrukturen är konstant. Sätt

$$X = \text{antal återanvända flaskor}$$

det gäller då att $X \in Bin\,(18, 0.9)$. Av detta följer att

a. $P\,(X \geq 15) = \sum_{k=15}^{18} \binom{18}{k} 0.9^k 0.1^{18-k} = 0.901\,8$

b. $\mu_X = 18 \times 0.9 = 16.2$ och $\sigma_X = \sqrt{18 \times 0.1 \times 0.9} = 1.2728$

Lösning 2: Sätt

$$\begin{aligned} X_{Ai} &= \text{antal hela glas med den nye leverantören} \\ X_{Bi} &= \text{antal hela glas med den gamle leverantören} \end{aligned}$$

Vi har två små oberoende stickprov med $n_A = 12$ och $n_B = 18$ och för dessa finner vi

$$\begin{aligned} \bar{X}_A &= \frac{1085}{12} = 90.42 \\ \bar{X}_B &= \frac{1528}{18} = 84.89 \end{aligned}$$

samt

$$\begin{aligned} s_A^2 &= \frac{1}{11}\left(98191 - 12\left(\frac{1085}{12}\right)^2\right) = 8.08 \\ s_B^2 &= \frac{1}{17}\left(129898 - 18\left(\frac{1528}{18}\right)^2\right) = 11.05 \end{aligned}$$

Konfidensintervallet kan nu skrivas

$$\begin{aligned} \bar{X}_A - \bar{X}_B \pm t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}} &= 5.53 \pm 2.048\sqrt{\frac{8.08}{12} + \frac{11.05}{18}} \\ &= (5.48 - 2.32, 5.48 + 2.32) \\ &= (3.16, 7.8) \end{aligned}$$

Således gäller med 95% säkerhet att $\mu_A \neq \mu_B$ varför den gamla leverantören ersätts med den nya eftersom intervallet ligger på den positiva sidan.

Lösning 3: De fyra hypotesstegen är:

a. Vi utgår ifrån att serien är bättre ty om detta är sant vill vi ha kontroll över sannolikheten att förkasta denna hypotes

$$H_0 : \pi \leq 0.10 \quad H_A : \pi > 0.10$$

b. Vi har ett stort stickprov varför testvariabeln blir

$$z = \frac{P - \pi_0}{\sqrt{\frac{P(1-P)}{n}}}$$

c. Det kritiska värdet är $z_{0.05} = 1.64$ (ensidigt test)

d. Observerat värde blir

$$z_{obs} = \frac{\frac{22}{198} - 0.1}{\sqrt{\frac{\frac{22}{198}\left(1 - \frac{22}{198}\right)}{198}}} = 0.497$$

Eftersom detta värde är mindre än 1.64 förkastas ej nollhypotesen på signifikansnivån 5%. Den nya serien är sannolikt bättre.

Tillämpad statistik

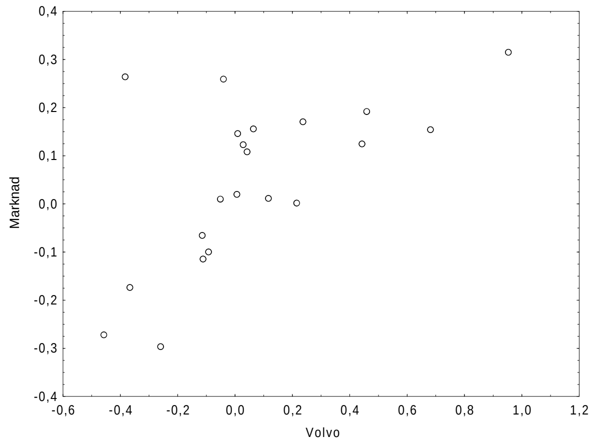
Lösning 4: Vi börjar med att ange den fullständiga variansanalystabellen

Source	SS	df	MS	F
Regression	0.902587	1	0.902587	11.39709
Error	1.504696	19	0.079195	
Total	2.407283	20		

samt beräknar

$$\sum (x_i - \bar{x})^2 = 2.497326 - 21 * 0.065481^2 = 2.40728.$$

a. (1)



(2) Följande blir de fyra stegen

Steg 1 $H_0 : \beta = 0 \quad H_1 : \beta \neq 0$

Steg 2 Som testvariabel tar vi

$$F = \frac{\text{Regression}/1}{\text{Error}/19}$$

där $F \in F(1, 19)$.

Steg 3 Beslutsregeln erhålls ur ekvationen

$$\begin{aligned} \alpha &= 0.05 = P(\text{förkasta } H_0 \text{ givet } H_0 \text{ sann}) \\ &= P(F > a) \\ a &= F_{0.05}(1, 19) = 4.38 \end{aligned}$$

Steg 4 Eftersom $F_{obs} = 11.4 > 4.38$ förkastas nollhypotesen.

Linjär regression är signifikant.

b. För förklaringsgraden gäller

$$R^2 = \frac{\text{Regression}}{\text{Total}} = \frac{0.902587}{2.407283} = 0.37494$$

varför $R \approx \sqrt{0.37494} \approx 0.61$ d v s 61 procent av variationen förklaras av en linjär modell.

c. Som prognos hittar vi

$$\hat{y} = 0.004370 + 1.244511 * 0.12 = 0.153711$$

och prediktionsintervallet blir

$$0.153711 \pm 2.093 * \sqrt{0.079195}\sqrt{1 + \frac{1}{21} + \frac{(0.12 - 0.065481)^2}{2.40728}}$$

d v s $(-0.449509, 0.756931)$.

Lösning 5: De fyra termerna står för

$$\begin{aligned} \mathrm{TR}_t &= \text{seriens trendkomponent}, \\ \mathrm{SN}_t &= \text{seriens säsongkomponent}, \\ \mathrm{CL}_t &= \text{seriens konjunkturkomponent}, \\ \mathrm{IR}_t &= \text{seriens irreguljära komponent}. \end{aligned}$$

a. Perioden släcks genom att addera fyra på varandra följande värden d v s man bildar serien

$$\bar{y}_t = \frac{y_t + y_{t+1} + y_{t+2} + y_{t+3}}{4}.$$

Denna nya serie blir dock förskjuten i förhållande till den ursprungliga, med ett halvt tidssteg, och man bildar därför istället

$$\bar{y}_t = \frac{y_t + 2y_{t+1} + 2y_{t+2} + 2y_{t+3} + y_{t+4}}{8}.$$

b. När en linjär trend föreligger ansätts modellen

$$\mathrm{tr}_t = \beta_0 + \beta_1 t.$$

c. Säsongskomponenten bestäms genom att man betraktar differensen

$$\mathrm{sn}_t = y_t - \bar{y}_t$$

d. Konjunkturen erhålls om man först bildar differenserna

$$\mathrm{cl}_t + \mathrm{ir}_t = y_t - \mathrm{tr}_t - \mathrm{sn}_t$$

och därefter addderar säg tre värden

$$\overline{\mathrm{cl}}_t = \frac{\mathrm{cl}_t + \mathrm{ir}_t + \mathrm{cl}_{t+1} + \mathrm{ir}_{t+1} + \mathrm{cl}_{t+2} + \mathrm{ir}_{t+2}}{3}.$$

Lösning 6: Följande gäller

a. En pessimist håller sig till Minimax eller **Maximin** kriterierna:

	Marknaden			Rad
Värdepapper	G	D	S	min
A	2000	**1200**	1500	1200
B	3000	**800**	1000	800
C	2500	**1000**	1800	1000
	Kolumn max			1200

Välj värdepapper A.

b. En optimist håller sig till Minimin eller **Maximax** kriterierna:

	Marknaden			Rad
Värdepapper	G	D	S	max
A	**2000**	1200	1500	2000
B	**3000**	800	1000	3000
C	**2500**	1000	1800	2500
	Kolumn max			3000

Välj värdepapper B.

c. Eftersom de olika framtiderna har samma sannolikhet att inträffa tilldelar vi dem sannolikheten $\frac{1}{3}$ detta ger

$$\begin{aligned}
E(A) &= 2000 \times \frac{1}{3} + 1200 \times \frac{1}{3} + 1500 \times \frac{1}{3} = 1566 \\
E(B) &= 3000 \times \frac{1}{3} + 800 \times \frac{1}{3} + 1000 \times \frac{1}{3} = 1600 \\
E(C) &= 2500 \times \frac{1}{3} + 1000 \times \frac{1}{3} + 1800 \times \frac{1}{3} = 1766
\end{aligned}$$

Välj värdepapper C.

d. Först beräknar vi besvikelsetabellen

	Marknaden			Rad
Värdepapper	G	D	S	max
A	**1000**	0	300	1000
B	0	400	**800**	800
C	**500**	200	0	500
	Kolumn min			500

Välj värdepapper C.