



## Försättsblad till skriftlig tentamen vid Linköpings Universitet

<b>Datum för tentamen</b>	2013-03-23
<b>Sal (1)</b> Om tentan går i flera salar ska du bifoga ett försättsblad till varje sal och <u>ringa in</u> vilken sal som avses	TER1
<b>Tid</b>	8-13
<b>Kurskod</b>	732G21
<b>Provkod</b>	TEN1
<b>Kursnamn/benämning</b> <b>Provnamn/benämning</b>	Sambandsmodeller Tentamen
<b>Institution</b>	IDA
<b>Antal uppgifter som ingår i tentamen</b>	4
<b>Jour/Kursansvarig</b> Ange vem som besöker salen	Lotta Hallberg
<b>Telefon under skrivtiden</b>	
<b>Besöker salen ca kl.</b>	10
<b>Kursadministratör/kontaktperson</b> (namn + tfnr + mailaddress)	annelie.almquist@liu.se, 2934
<b>Tillåtna hjälpmedel</b>	Räknedosa. Läroboken: <i>Applied linear statistical models</i> av Kutner, Nachtsheim m fl som inte får innehålla anteckningar men får ha markeringar och flärpar. Flärpar får ha en liten anteckning.
<b>Övrigt</b>	G=12, VG=16

## TENTAMEN I SAMBANDSMODELLER, 2013-03-23

**Skrivtid:** kl: 8-13

**Hjälpmedel:** Räknedosa. Läroboken: *Applied linear statistical models* av Kutner, Nachtsheim m fl som inte får innehålla anteckningar men får ha markeringar och flärpar. Flärpar får ha en liten anteckning.

**Jourhavande lärare:** Lotta Hallberg.

Redovisa och motivera kort alla dina lösningar

---

Tolka (om möjligt) alla dina resultat!

1

Vid produktion av 32 delar av cellulosa från bomull användes olika inställningar på koncentrationsnivå av formaldehyd (Conc), katalysatorkvot (Ratio) och härdnings-temperatur (Temp). Ett mått på hållbarheten då produkten pressats kallas här rating (Rating), vilket är ett mått på skrynkelhårdigheten. Rating är responsvariabel.

Nedan visas ett utdrag från de 32 delarna.

Conc	Ratio	Temp	Rating
8	4	100	1,40000
2	4	180	2,20000
7	4	180	4,60000
10	7	120	4,90000
.	.	.	.
7	10	100	4,60000
4	1	100	1,40000
4	1	100	1,45000

### Korrelationsmatris

	Conc	Ratio	Temp
Ratio	0,112		
Temp	0,154	0,116	
Rating	0,369	0,748	0,365

### Modell

#### Regression Analysis: Rating versus Conc; Ratio; Temp

The regression equation is

$$\text{Rating} = -0,447 + 0,157 \text{ Conc} + 0,215 \text{ Ratio} + 0,0111 \text{ Temp}$$

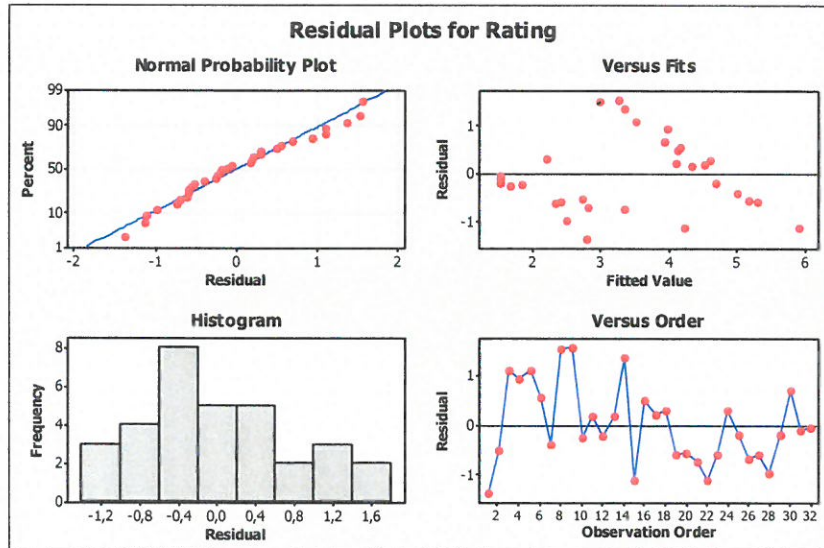
Predictor	Coef	SE Coef	T	P
Constant	-0,4468	0,7367	-0,61	0,549
Conc	0,15718	0,06531	2,41	0,023
Ratio	0,21537	0,03262	6,60	0,000
Temp	0,011097	0,004761	2,33	0,027

$S = 0,837165$     $R\text{-Sq} = 70,0\%$     $R\text{-Sq}(\text{adj}) = 66,8\%$

$\text{PRESS} = 25,1293$     $R\text{-Sq}(\text{pred}) = 61,59\%$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	45,796	15,265	21,78	0,000
Residual Error	28	19,624	0,701		
Total	31	65,420			



- En modell är anpassad ovan. Utför residulanalys för denna modell. 1p
- I utskriften är PRESS beräknat. Tolka resultatet. 1p
- Hur har  $R\text{-Sq}(\text{pred}) = 61,59\%$  beräknats? 1p
- Prediktera Rating då  $\text{Conc}=8$ ,  $\text{Ratio}=10$  och  $\text{Temp}=140$ . Beräkna även ett 95% konfidensintervall. Medelfelet för  $\hat{Y}$  är 0.225. 2p
- Vilken variabel ska man börja med vid framåtvalsmetoden för detta datamaterial? Förklara. 1p
- Hur mycket ökar Rating i snitt då Temp ökar med 20 grader? 1p

## 2

Ett företag som tillverkar frukostflingor vill bland annat undersöka vilken effekt en TV-reklamkampanj har haft på försäljningen på en ny produkt kallad Cocoa Crunch. Efter att ha visat reklamen i en vecka i ett visst område så slumpades 71 vuxna ut som gick ut ur en viss stormarknad i området. Man ställde bland andra frågorna:

- har du köpt ett Cocoa Crunch paket (Bought)
- har du barn (Children)
- har du sett reklamen (ViewAd)

Följande modell anpassades:

### Binary Logistic Regression: Bought versus Children; ViewAd

Link Function: Logit

#### Response Information

Variable	Value	Count	
Bought	1	22	(Event)
	0	49	
	Total	71	

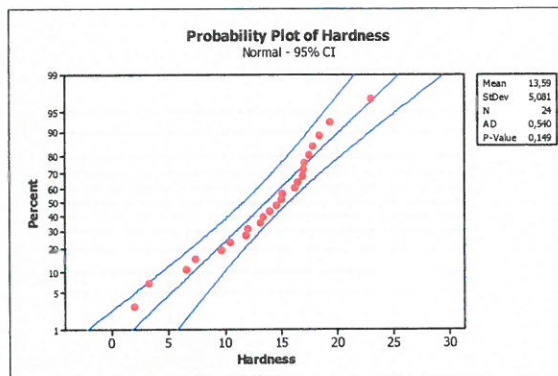
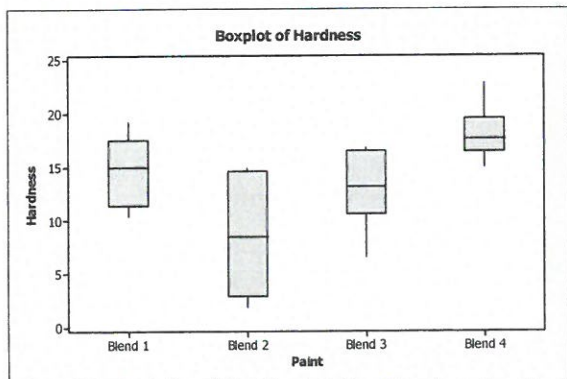
#### Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Constant	-2,70017	0,819416	-3,30	0,001			
Children							
Yes	1,64148	0,812597	2,02	0,043	5,16	1,05	25,38
ViewAd							
Yes	1,10588	0,562370	1,97	0,049	3,02	1,00	9,10

- Skriv upp och visa hur den skattade modellen ser ut. 1p
- Tolka de båda oddskvoterna. 2p
- Prediktera sannolikheten att köpa ett paket Cocoa Crunch då familjen har barn och då de sett annonsen 1p

## 3

I en undersökning så jämfördes hårdheten hos fyra olika blandningar av färg (blend). Sex prov med färg applicerades på små metallplattor. Färgen härdades, sedan mättes hårdheten (Hardness).



### Descriptive Statistics: Hardness

Variable	Paint	Mean	StDev
Hardness	Blend 1	14,73	3,36
	Blend 2	8,57	5,50
	Blend 3	12,98	3,73
	Blend 4	18,07	2,64

### One-way ANOVA: Hardness versus Paint

Source	DF	SS	MS	F
Paint	*	*	93,9	*
Error	*	*	15,6	
Total	*	593,8		

- Beskriv modellen som har anpassats och avgör om modellantagandena kan anses vara uppfyllda? 1p
- Fyll i ANOVA-tabellen ovan och testa om det är skillnad mellan de olika blandningarna. Var nog med hypoteserna. 2p
- Beräkna alla Tukey - intervall och tolka resultatet. 2p

## 4

Forts uppgift 3.

Anta nu att de fyra färgblandningarna valdes slumpmässigt från väldigt många blandningar.

- Beräkna ett 95% konfidensintervall för  $\sigma_{\mu}^2$ . 1,5p
- Beräkna ett 95% konfidensintervall för  $\sigma_{\mu}^2/\sigma_Y^2$ . 2,5p