

Háskóli Íslands	<b>09.10.16 Línuleg algebra og tölfraði</b>	Raunvísindadeild
Föstuudagur	<b>12. desember 2003</b>	<b>kl 13:30-16:30</b>
<b>Leyfileg hjálpargögn:</b> Glósur, bækur og reiknivélar	<b>Athugið</b> að fartölvur eru ekki leyfðar.	<b>Vægi</b> dæma er gefið: 100 stig teljast full lausn, en alls eru stigin fleiri.

Notið 5% marktækniröfu nema annað sé tekið fram. Munið að taka skýrt fram núlltilgátur og gagntilgátur þar sem það á við.

Eftirfarandi fylki  $\mathbf{X}$  og vigur  $\mathbf{y}$  verða notuð í dæmum 1-3:

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$

- (10)** Finnið þann vigur,  $\hat{\beta}$  sem lágmarkar  $\|\mathbf{y} - \mathbf{X}\hat{\beta}\|^2$ .
- (10)** Gefið dálkvigrum fylkisins  $\mathbf{X}$  nöfnin  $\mathbf{u}$ ,  $\mathbf{v}$  og  $\mathbf{w}$  og finnið hornréttan einingargrunn fyrir spann þeirra vigra,  $V = sp(\mathbf{u}, \mathbf{v}, \mathbf{w})$ .
- (5)** Finnið þann vigur,  $\hat{\mathbf{y}}$ , sem er næstur  $\mathbf{y}$  meðal allra vigra í  $V$ .
- (10)** Finnið eigingildi fylkisins

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

- (15)** Teiknið mynd og nýtið hana til að finna þau gildi,  $x$  og  $y$ , sem hámarka  $z = y - x$  með tilliti til

$$\begin{aligned} x + 2y &\geq 2 \\ x &\leq 2 \\ 4x + 3y &\leq 12 \\ 5x - 3y &\geq -3 \\ x &\geq 0 \\ y &\geq 0 \end{aligned}$$

Munið að rökstyðja svarið.

- (10)** Gerðar voru mælingar á hæð trjáa ( $y_{ij}$ ) á 2 mismunandi stöðum ( $i$ ) á landinu, 3 mælingar ( $j$ ) á hverjum stað. Táknið staðaráhrif sem frávik frá meðaltali ( $\mu$ ) með  $\alpha_i$  að viðbættum áhrifum vegna hitastigs (mælt sem "gráðudagar"),  $x_{ij}$ .

Setjið nákvæmlega fram á fylkjaformi líkan fyrir samvikagreiningu þar sem hæð trjáa er spáð með staðaráhrifum og hitaáhrifum. Notið tölur eða tákni eftir því sem við á til að sýna öll stök í öllum vigrum eða fylkjum.

**7. (20)** Mælingar eru gerðar á einstaklingum sem hafa orðið fyrir áfalli. Mælingar fyrir hvern einstakling eru teknar saman í staðlaða heildareinkunn fyrir ásigkomulag og mynda mælikvarða á úthald eða þol. Úrtak 5 slíkra einstaklinga gaf mælingarnar 85, 94, 97, 89, 101.

(a) Heilbrigður einstaklingur fær að meðaltali einkunnina 100, sem er viðmiðun í kerfinu. Er hægt að fullyrða að áfallið leiði til minnkunar á úthaldi?

(b) Eftir 5 mánaða endurhæfingu voru sömu einstaklingar mældir aftur og fengust þá niðurstöðurnar 84, 96, 102, 88, 105. Hefur endurhæfingin haft áhrif?

(c) Stærri könnun meðal 49 einstaklinga leiddi í ljós að 38 fengu hærri einkunn eftir endurhæfingu en fyrir. Er þetta marktæk breyting?

**8. (10)** Teiknið fall sem vex línulega frá  $f(0) = 0$  upp í  $f(1) = k$  og minnkar síðan línulega úr  $f(1) = k$  í  $f(3) = 0$ . Utan bilsins  $(0, 4)$  er  $f(x) = 0$ .

(a) Hvað þarf  $k$  að vera til að  $f$  sé þéttifall?

(b) Ef  $X$  er hending með þéttifallið  $f$ , hvað er þá  $P[X \geq 2]$ ?

**9. (10)** Á kennsluvefnum voru tekin út fyrsta og síðasta svar hvers nemanda við spurningu úr hverjum fyrirlestri. Fyrir hverja spurningu fæst svarið 0 eða 1. Nemendur fengu að meðaltali einkunnina 0.53 við fyrstu spurningu í hverjum fyrirlestri en 0.77 fyrir þá síðustu. Meðaltölin eru byggð á 1539 svörum, hvort fyrir sig.

(a) Prófið hvort munurinn er marktækur og túlkið niðurstöðuna?

(b) Hverjar eru forsendur prófsins og standast þær?

10. (30) Kanna skal hvernig mælingar í vigrinum  $\mathbf{y}$  tengist stýribreytum í vigrunum  $\mathbf{x}$ ,  $\mathbf{w}$  og  $\mathbf{z}$ . Sérstakur áhugi er á að kanna tengsl  $\mathbf{y}$  við allar breytur og að kanna hvort lýsa megi  $\mathbf{y}$  með  $\mathbf{x}$  einu og sér.

Til að prófa þetta eru gefnar tvær R skiparnir, annars vegar

**fmF<-lm(y~w+z+x)**

til að meta  $y_i = \alpha + \beta w_i + \gamma z_i + \delta x_i + \epsilon_i$  (stórt líkan) og hins vegar

**fmR<-lm(y~x)**

til að meta  $y_i = \alpha + \delta x_i + \epsilon_i$  (minna líkan). Gert er ráð fyrir óháðum normaldreifðum frávikum.

Líkönin eru síðan greind með nokkrum R skipunum, eins og fram kemur á næstu síðu.

Munið að rökstyðja svörin með tilvísun í tiltekna(r) tölu(r) í úttakinu.

- (a) Er nauðsynlegt að hafa  $x$  í lokallíkaninu?
- (b) Er nauðsynlegt að hafa  $z$  í lokallíkaninu?
- (c) Hvert er matið á  $\delta$  í stóra líkaninu?
- (d) Hvað er  $MSE$  í stóra líkaninu?
- (e) Hvað er  $SSTOT$ ?
- (f) Hvert er óvissumatið á  $\gamma$  (þ.e.  $\hat{\sigma}_\gamma$ )?
- (g) Er nóg að spá  $y$  með  $x$  eingöngu?
- (h) Hvað er  $SSE$  í litla líkaninu?

```
> anova(fmF)
```

```
Analysis of Variance Table
```

```
Response: y
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
w	1	3269.6	3269.6	1270.204	< 2.2e-16 ***
z	1	55.7	55.7	21.654	0.0001364 ***
x	1	39.0	39.0	15.148	0.0008407 ***
Residuals	21	54.1	2.6		

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> anova(fmR, fmF)
```

```
Analysis of Variance Table
```

```
Model 1: y ~ x
```

```
Model 2: y ~ w + z + x
```

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	23	214.016				
2	21	54.056	2	159.960	31.071	5.311e-07 ***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> drop1(fmF, test="F")
```

```
Single term deletions
```

```
Model:
```

```
y ~ w + z + x
```

	Df	Sum of Sq	RSS	AIC	F value	Pr(F)
<none>			54.056	27.279		
w	1	156.054	210.110	59.219	60.6250	1.268e-07 ***
z	1	3.906	57.962	27.023	1.5173	0.2316384
x	1	38.993	93.048	38.856	15.1481	0.0008407 ***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> summary(fmF)
```

```
Call:
```

```
lm(formula = y ~ w + z + x)
```

```
Residuals:
```

Min	1Q	Median	3Q	Max
-2.4341	-1.4062	0.2688	1.1839	2.6240

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4.3381	1.6857	2.574	0.01771 *
w	6.8344	0.8778	7.786	1.27e-07 ***
z	-1.4251	1.1569	-1.232	0.23164
x	0.9825	0.2524	3.892	0.00084 ***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.604 on 21 degrees of freedom
```

```
Multiple R-Squared: 0.9842,    Adjusted R-squared: 0.9819
```

```
F-statistic: 435.7 on 3 and 21 DF,  p-value: < 2.2e-16
```