

Eksamen i STAT100- statistikk grunnkurs

Bokmål

Tirsdag 24. mai 2005

C2, alle trykte og skrevne hjelpemidler, samt lommekalkulator.

Ved vurdering teller alle delpunkter (1a, 1b etc) likt.

Oppgaveteksten er på 3 sider

Kontakt under eksamen: Trygve Almøy, tlf: 64 96 58 20

Oppgave 1

En bonde dyrker to forskjellige byggsorter (A og B). For begge sortene er avling pr. dekar normalfordelt. Lang erfaring har vist at begge sortene har forventet avling på 500 kg og et standardavvik på 30 kg for sort A og et standardavvik på 50 kg for sort B.

a) Bonden påstår at sort A gir mer stabil avling enn sort B. Er du enig med han i dette? Det regnes som et dårlig år dersom avlingen er under 440 kg. Hva er sannsynligheten for et dårlig år ved sort A? Er det mer sannsynlig med dårlig år ved sort B? (Grunngi svaret.)

b) De 10 % beste årene regnes som kronår. Hvor stor avling med sort A må det være for at dette skal skje? Hva er sannsynligheten for at 2 av 3 påfølgende år regnes som kronår (la avling hvert år være uavhengig av det foregående).

Oppgave 2

I en undersøkelse (beskrevet i *New England Journal of Medicine*, 1997) ble 615 røykere som alle ville slutte å røyke utsatt for følgende: 309 fikk (etter loddtrekning) Zyban, et legemiddel som motvirker depresjon, resten fikk bare en harmløs sukkerpille (såkalt placebo). Medisinen ble gitt i seks uker. Ingen fikk beskjed om hva slags medisin de fikk. Merk at du kan gå rett til oppgave b dersom du ikke klarer oppgave a

a) Et år etter hadde 71 av de som fikk Zyban sluttet å røyke, mens 37 av de som fikk sukkerpillen hadde sluttet. Vil du på bakgrunn av dette påstå at antidepressiva helt sikkert er effektivt som hjelpemiddel ved røykeslutt?

Siden alle var motivert til å slutte, men ikke klarte det helt, ville en også studere om bruken av antidepressiva førte til reduksjon i sigarettforbruket blant de som ikke sluttet. Alle ble bedt om å rapportere gjennomsnittlig daglig sigarettforbruk.

b) Estimer forventet reduksjon i daglig sigarettforbruk ved bruk av antidepressiva. Finn et 95% konfidensintervall for forventet reduksjon, og gi en tolkning av intervallet. Dette intervallet dekker ikke 0. Hva kan du si om effekten av antidepressiva ved røykeslutt? (*Bruk utskriften nedenfor så mye du kan*)

	N	Mean	StDev	SE Mean
placebo	269	14,39	3,67	0,22
zyban	238	13,69	3,11	0,20

Difference = $\mu(p) - \mu(z)$

Estimate for difference: 0,692790

95% CI for difference: (0,094441; 1,291139)

Both use Pooled StDev = 3,4224

En ville også se på sammenhengen mellom nikotin (N) og karbonmonoxid (CO) i sigaretter begge deler målt i milligram. En valgte ut 203 forskjellige sigarettmerker og målte CO og N. Et spredningsplot (scatterplot) finner du på slutten av denne oppgaven. Data ble analysert med en enkel lineær regresjonsanalysemodell:

$$CO_i = \beta_0 + \beta_1 N_i + \varepsilon_i, \text{ der } \varepsilon_i \text{ er } N(0, \sigma) \text{ og uavhengige. } i = 1, 2, 3 \dots, 203.$$

Utskrift av analysen finner du her:

Regression Analysis: co versus nikotin

The regression equation is
 $CO = 1,04 + 12,1 * N$

Predictor	Coef	SE Coef	T	P
Constant	1,0393	0,4632	2,24	0,026
nikotin	12,1158	0,5149	23,53	0,000

S = 2,78840

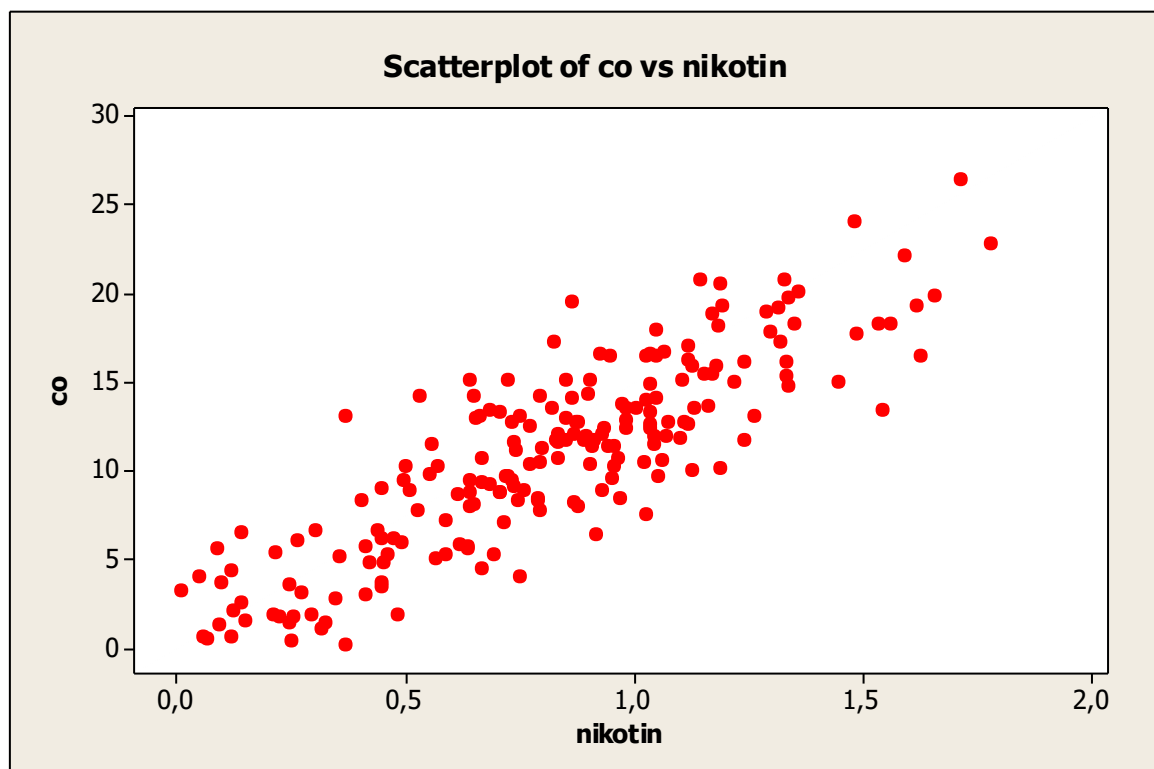
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	4305,4	4305,4	553,73	0,000
Residual Error	201	1562,8	7,8		
Total	202	5868,2			

Gjennomsnittlig nikotininhold er 0,81 milligram

c) Tolk og estimer alle parametere i modellen. Finn også R^2 og gi denne en tolkning.

d) Prediker CO-innholdet i en sigarett med nikotininhold på ett milligram. Finn også et 95 % prediksjonsintervall for denne verdien.



Oppgave 3

Volum (målt i milliliter) av venstre hjertekammer ble målt på et tilfeldig utvalg mannlige toppidrettsutøvere innen svømming, langdistanseløping, langrenn og bryting. Resultatet er vist på slutten av denne oppgaven.

Data ble analysert ved enveisvariansanalyse modell:

a) Sett opp modellen. Fyll inn manglende verdier i tabellen nedenfor. Kan du påvise forskjell i forventet hjertevolum mellom de forskjellige idrettene?

One-way ANOVA

Source	DF	SS	MS	F
Factor	*	5579	*	*
Error	*	*	*	
Total	*	9890		

Sp = 10,94

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
svømming	10	174,73	10,19	(-----*-----)
løping	10	164,07	12,00	(-----*-----)
langrenn	10	187,74	6,21	(-----*-----)
bryting	10	156,30	13,88	(-----*-----)

b) Test om det er forskjell i forventet hjertevolum mellom langrennsløpere og de andre idrettene.

Data for oppgave 3

Row	svømming	løping	langrenn	bryting
1	177	180	193	166
2	178	152	185	153
3	177	182	192	164
4	162	157	191	132
5	158	171	186	152
6	184	172	183	151
7	194	164	197	182
8	171	145	192	166
9	176	162	181	152
10	171	157	178	145