

Nynorsk

Fagleg kontakt under eksamen:

John Tyssedal

Telefon 41645376

## EKSAMEN I EMNE TMA4240/TMA4245 STATISTIKK

5. august 2009

Tid: 15:00-19:00

Hjelpemiddel: *Tabeller og formler i statistikk*, Tapir Forlag  
K. Rottmann: *Matematisk formelsamling*  
Kalkulator HP30S/CITIZEN SR-270X  
Gult, stempla A5-ark med egne handskrivne notat.

Sensuren fell: 26. august 2009

### Oppgåve 1 Klubbmesterskapet

Det skal arrangerast klubbmesterskap i langrenn i sportsklubben Dalen IL. I seniorklassen stiller berre to løparar; Asbjørn og Bernt. For å auke spenninga har klubben bestemt at oppmannen skal preparere to par ski,  $S_1$  og  $S_2$ , på ulik måte, og la løparane trekkje ski tilfeldig.

Grunna variasjon i føreforhold og dagsform går ein ut i frå at tida løparane brukar på løypa er ein tilfeldig variabel. Ein går vidare ut i frå at gitt kva skipar løparane brukar, så vil tidene til løparane vere uavhengige. Dersom Asbjørn brukar skipar  $S_1$  vil tida vere Gaussisk (normal) fordelt med forventning 60 minutt og standardavvik 5 minutt. Da må Bernt ha skipar  $S_2$ , og tida hans vil vere Gaussisk (normal) fordelt med forventning 65 minutt og standardavvik 5 minutt.

- a) Kva er sannsynet for at Asbjørn brukar mindre enn 55 minutt på løypa gitt at han går på skipar  $S_1$ ?

Kva er sannsynet for at Asbjørn brukar mellom 55 og 60 minutt på løypa gitt at han går på skipar  $S_1$ ?

- b) Kva er sannsynet for at Asbjørn på skipar  $S1$  brukar mindre tid enn Bernt på skipar  $S2$ ?

Dersom Asbjørn brukar skipar  $S2$  går ein ut i frå at tida er Gaussisk (normal) fordelt med forventning 63 minutt og standardavvik 7 minutt. Bernt får då skipar  $S1$  og ein går ut i frå at tida er Gaussisk (normal) fordelt med forventning 62 minutt og standardavvik 4 minutt.

- c) Gå ut i frå at skipara  $S1$  og  $S2$  blir gitt tilfeldig med same sannsyn til Asbjørn og Bernt.

Gitt at Asbjørn brukte mindre tid enn Bernt, kva er sannsynet for at Asbjørn gjekk på skipar  $S1$ ?

## Oppgåve 2 Telefonseljaren

En student har sommarjobb som telefonseljar. Ein går ut i frå at tida han brukar på kvar samtale,  $T$ , er eksponential fordelt med parameter  $\beta$ , som er avhengig av produktet han skal selje:

$$T \sim e(t; \beta) = \begin{cases} 0 & ; t < 0 \\ \frac{1}{\beta} \exp\{-\frac{t}{\beta}\} & ; t \geq 0 \end{cases}$$

Videre går vi ut i frå at samtalene følgjer umiddelbart etter kvarandre utan pause, med tider som er uavhengige.

For første produkt han sel er  $\beta = 4.0$  minutt, som skal brukast i a) og b).

- a) Kva er sannsynet for at ei vilkårleg samtale varer lenger enn 8 minutt?

Kva er sannsynet for at minst ein av to påfølgjande samtaler varer lenger enn 8 minutt?

- b) En dag har studenten 10 samtaler. Kva er sannsynet for at to eller færre av desse samtalene varer lenger enn 8 minutt?

Etter to veker på jobb skal studenten selje eit nytt produkt ved telefonsal. Han ynskjer å berekne parameteren  $\beta$  for dette produktet, og registrerer lengda på fem samtaler:

Samtale nr $i$ :	1	2	3	4	5
Lengde $t_i$ (min):	3.8	3.5	2.9	5.3	6.0

- c) Utlei sannsynsmaksimeringsestimatorene for  $\beta$  basert på observasjonene over.

Kva blir estimatet for  $\beta$ .

Dersom samtaletidene er eksponentielt fordelte og uavhengige er akkumulert tid for  $n$  samtaler,  $W_n = \sum_{i=1}^n T_i$ , gammafordelt med parametere  $(\beta, n)$ :

$$W_n \sim g(w_n; \beta, n) = \begin{cases} 0 & ; w_n < 0 \\ \frac{1}{\beta^n n!} w_n^{n-1} \exp\{-\frac{w_n}{\beta}\} & ; w_n \geq 0 \end{cases}$$

Den andre veka med det nye produktet vel studenten å registrere talet på samtaler og akkumulert tid fram til lunsj hver dag:

Dag nr $i$ :	1	2	3	4	5
Antall samtaler $n_i$	11	10	13	9	10
Akkumulert tid $w_{n_i}$ (min):	45.3	39.0	55.3	40.2	50.2

- d) Utlei sannsynsmaksimeringsestimatorene for  $\beta$  basert på dei akkumulerte dagsobservasjonane over.

Undersøk om denne estimatoren og estimatoren fra punkt c) er forventingsrette. Samanlikn og variansen til dei to estimatorane.

Dersom ein ynskjer ein forventningsrett estimator med minst moglege varians, og berre kan samle inn fem data som i designa over, kva innsamlingsdesign bør ein velje?

### Oppgåve 3 Slankegruppa

Eit helsestudio har ei slankegruppe med 13 personar. For å kartlegge vekttafet ved sykling på ergometersykkel sykklar 12 av deltakarane ulik tid med samme hastighet, samt at ein veg dei før og etter. Sykkeltid og vekttafet blir registrert:

Deltakar nr $i$ :	1	2	3	4	5	6	7	8	9	10	11	12
Sykkeltid $t_i$ (min):	45.1	35.8	43.1	50.6	56.4	49.9	52.9	58.3	47.0	56.1	49.0	42.0
Vekttafet $\Delta v_i$ (kg):	1.04	0.98	1.03	1.19	1.29	1.33	1.39	1.53	1.32	1.43	1.14	1.06

Følgjande middelerdiar blir oppgitt:

$$\frac{1}{12} \sum_i t_i = 48.9 \quad \frac{1}{12} \sum_i \Delta v_i = 1.23 \quad \frac{1}{12} \sum_i \Delta v_i t_i = 60.9 \quad \frac{1}{12} \sum_i t_i^2 = 2426.9$$

Ein regresjonsmodell blir brukt for å finne samanhengen mellom vekttafet og sykkeltid:

$$\Delta V = \beta t + R$$

der  $\beta$  er en ukjent regresjonsparameter som skal bestemast, og  $R$  er ein Gaussisk (normal) fordelt tilfeldig variabel med forventning 0 og kjend varians  $\sigma^2 = 0.1^2$ .

I forsøket over går ein ut i frå at feilledda  $R$  er uavhengige.

- a) Gje ei tolkning av parameteren  $\beta$ .

Utlei ein optimal (sannsynsmaksimering eller minste kvadrat) estimator for  $\beta$ .

Rekn ut estimatet for  $\beta$ .

I dei neste punkta skal følgjande estimator for  $\beta$  brukast:

$$\hat{\beta} = \frac{\sum_{i=1}^{12} \Delta v_i t_i}{\sum_{j=1}^{12} t_j^2}$$

- b) Helsestudioet reklamerer med at forventa vekttao ved ergometersykling er minst 1.6 kg i timen.

Gjev dataane over deking for en slik påstand?

Formuler svaret som ein hypotesetest med signifikansnivå 0.05.

- c) Ein av deltakarane i gruppen var fråverande under testen. Ho planlegg å sykle 50 minutt på samme måte som dei andre.

Utlei eit 95% prediksjonsintervall for vekttaoet  $\Delta V$  ho kan rekne med å oppnå.