Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Nynorsk

Fagleg kontakt under eksamen:
John Tyssedal 73 59 35 34
Arvid Næss 99 53 83 50

SIF5060 Statistikk

Torsdag 28. november 2002
Tid: 09:00–14:00

Hjelpemiddel: Godkjent lommekalkulator med tomt minne.
Statistiske tabeller og formler, Tapir forlag.
K. Rottman: Matematisk formelsamling.

Sensuren fell 12. januar 2003.

Oppgave 1 Skader av jordskjelv

Ein by ligg i eit område som er utsett for jordskjelv. Det neste jordskjelvet som rammar området, kan klassifiserast som kraftig (K), moderat (M) eller svakt (S). Ein vilkårlig bygning av ein bestemt type vil svikte dersom akkumulert skade D blir større enn ein kritisk verdi $d_0 = 1.0$. Gå ut i frå at D for kvar av dei tre jordskjelvklassene kan modellerast som ein tilfeldig variabel med sannsynstettleik

$$f_D(d) = \begin{cases} \lambda e^{-\lambda d} & , \quad d \geq 0 \\ 0 & , \quad d < 0. \end{cases}$$

Parameteren $\lambda > 0$ har ein verdi bestemt av intensiteten til jordskjelvet. Ein har funne at dei tre jordskjelvklassene, K, M og S har tilhøyrande parameterverdiar $\lambda_K = 1.61$, $\lambda_M = 3.00$ og $\lambda_S = 4.60$.

a) Bestem for kvar av dei tre jordskjelvklassane sannsynet for at ein vilkårleg bygning av den bestemte typen skal svikte.

Det er estimert at neste jordskjelv som rammar området, vil vere kraftig, moderat eller svakt med sannsyn i gitt rekkjefølgje $\mathrm{P}(K) = 0.02$, $\mathrm{P}(M) = 0.20$ og $\mathrm{P}(S) = 0.78$.

b) Bestem sannsynet for at ein bygning av den bestemte typen sviktar i neste jordskjelv.

Kva er sannsynet for at jordskjelvet var moderat dersom bygningen sviktar?

Gå ut i frå at A og B representerer to bygningar av den bestemte typen. Det er estimert at dersom bygning A sviktar, så er sannsynet for at bygning B og sviktar lik 0.50 for K, 0.15 for M og 0.02 for S.

c) Bestem sannsynet for at begge bygningane sviktar i neste jordskjelv.

Kva er sannsynet for at jordskjelvet ikkje var kraftig dersom bygning A svikta etter eit jordskjelv medan bygning B ikkje svikta?

Oppgave 2 Norske jordskjelv

Den absolutte styrken av eit jordskjelv blir målt på Richter-skalaen, som er ein logaritmisk skala. Det sterkaste norske jordskjelvet som er registrert, er målt til 5.8 på denne skalaen. I punkt **2a**) og **2b**) skal vi spesielt ta for oss jordskjelv registrert på den norske kontinental-sokkelen utanfor Midt-Norge av en slik styrke at dei kan registrerast av mennesker. Basert på historiske data fram til 1980, er det rimelig å gå ut i frå at den absolutte styrken av desse er normalfordelt med forventing $\mu = 4.2$ og standardavvik $\sigma = 0.4$.

a) For at vesentlege skader skal skje, reknar ein at styrken på Richter-skalaen må overstige 5.4. Finn sannsynet for at den absolutte styrken av eit vilkårlig slikt jordskjelv utanfor Midt-Norge skal overstige 5.4.

Finn og ein verdi slik at det er berre 5% sannsynleg at den absolutte styrken av slike jordskjelv skal overstige k.

Siden 1980 er det registrert 9 jordskjelv av den gitte typen med følgjande absolutte styrke målt på Richter-skalaen: 4.8, 3.5, 3.8, 5.3, 3.8, 3.5, 3.5, 3.5, 4.0. Du kan gå ut i frå at desse målingane er realiseringar av 9 uavhengige identisk normalfordelte variable $Y_1, Y_2, \ldots, Y_9$. For desse tala er

$$\bar{y} = 3.967 \quad \text{og} \quad \sum_{i=1}^{9} (y_i - \bar{y})^2 = 3.40.$$

Variansen skal vi no gå ut i frå er ukjend.

b) Ein er interessert i å finne ut om forventa absolutt styrke av slike jordskjelv har auka som ei følgje av den norske oljeaktiviteten. Formuler dette som ein hypotesetest med nullhypotese og alternativ hypotese.

Utfør testen. Kva blir konklusjonen når signifikansnivået er sett til 5%?

Det er og av interesse å finne ut om frekvensen av jordskjelv har forandra seg som følgje av oljeaktiviteten. Vi skal no ta for oss alle typar jordskjelv registrert over heile landet og gå ut i frå at talet på jordskjelv, X i eit tidsrom $[0, t]$ målt i år er Poissonfordelt med parameter λt , det vil sei

$$\mathrm{P}(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}.$$

c) Gå i dette punktet ut i frå at $\lambda = 5$.

Finn sannsynet for at det skal bli registrert 8 jordskjelv eit år.

Finn og sannsynet for at det skal bli registrert 5 eller fleire jordskjelv i eit tidsrom på eit halvt år.

Kva er sannsynet for at det største talet på jordskjelv pr. år over ein 10 års periode skal overstige 10?

La T vere tida frå ein tek til å registrere jordskjelv til første jordskjelv inntreff.

d) Vis at T er eksponentialfordelt med parameter $1/\lambda$.

Vis og at $2\lambda T$ er kjikvadrat-fordelt med 2 fridomsgrader.

e) Tida mellom to påfølgjande jordskjelv har same fordeling som T. La $T_1, T_2, \ldots, T_n$ vere tider mellom to påfølgjande jordskjelv og $t_1, t_2, \ldots, t_n$ være dei tilhøyrande registrerte tidene. Finn sannsynsmaksimeringsestimatoren for λ, $\widehat{\lambda}$, basert på det tilfeldige utvalet.

Vis korleis ein kan finne eit 95% konfidensintervall for λ basert på $2\lambda n/\widehat{\lambda}$.

Kva blir konfidensintervallet når gjennomsnittet av 10 tider mellom påfølgjande jordskjelv er 3 månader? Er det grunn til å konkludere med at $\lambda \neq 5$? Grunngjev svaret.

Oppgave 3 Automatisert laboratorium

I eit laboratorium ynskjer ein å evaluere samanhengen mellom to variablar Y og x. Apparaturen er sett opp slik at ein kan fastsetje x for deretter å måle Y. Ein vel å nytte følgjande modell for samanhengen mellom variablane

$$Y = \alpha + \beta x + \varepsilon$$

der α og β er to ukjende koeffisientar og ε er ein tilfeldig variabel som er normalfordelt med forventing 0 og ukjend varians σ^2. La $x_1, x_2, \ldots, x_n$ vere n verdiar av variabelen x og $y_1, y_2, \ldots, y_n$ dei tilhøyrande verdiane som blir målt for Y. Desse skal sjåast på som realiseringar av n uavhengige variablar $Y_1, Y_2, \ldots, Y_n$. Minste kvadratsums (least squares) estimatorane, A og B, for koeffisientane α og β er då gitt ved

$$A = \frac{1}{n}\sum_{i=1}^{n} Y_i - B\bar{x}$$

$$B = \frac{\sum_{i=1}^{n}(x_i - \bar{x})Y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

der

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

a) Vis at estimatorane A og B er forventingsrette estimatorar for α og β.

Bruk at kovariansen mellom $\bar{Y} = (1/n)\sum_{i=1}^{n} Y_i$ og B er 0 og utlei variansen til estimatorane A og B.

Kva sannsynsfordelingar har estimatorane A og B? Grunngjev svaret.

Laboratorieforsøka er svært arbeidskrevjande, men apparaturen er automatisert slik at forsøka kan utførast automatisk for ekvidistante verdiar av x. Sjå på to måleseriar med $n = 10$

Serie 1:	$x_1 = 1, x_2 = 2, \ldots, x_{10} = 10$
Serie 2:	$x_1 = 2, x_2 = 4, \ldots, x_{20} = 20$

Målet med forsøket er å predikterere Y_0 for $x_0 = 5.5$. Følgjande prediktor blir brukt: $\widehat{Y}_0 = A + Bx_0$.

b) Utlei variansen til $Y_0 - \widehat{Y}_0$.

Kva måleserie bør nyttast for å prediktere Y_0 best mogeleg for $x_0 = 5.5$? Grunngjev og kommenter svaret du har funne.