Noregs teknisk-naturvitskaplege universitet
Institutt for matematiske fag

Nynorsk

Fagleg kontakt under eksamen:

Ingelin Steinsland	926 63 096
Ola Diserud	932 18 823
Arvid Næss	995 38 350

TMA4245 Statistikk

Laurdag 26. mai 2012 kl. 9–13

Hjelpemiddel: Gult A5-ark med eigne handskrivne notat (stempla av Institutt for matematiske fag), *Tabeller og formler i statistikk* (Tapir forlag), *Matematisk formelsamling* (K. Rottmann), kalkulator HP 30s eller Citizen SR-270X

Sensur: 18. juni 2012

Oppgåve 1   Snødensitet

Densiteten (massetettleiken) av snø er høgst 1 kg/dm^3 (veldig våt snø). Gå ut frå at sannsynstettleiken for densiteten i kg/dm^3 av ei tilfeldig vald snøprøve er gitt ved $f(x) = \beta(\beta+1)x(1-x)^{\beta-1},\ 0 \leq x \leq 1$, der β er ein positiv parameter.

a) Gå ut frå, berre i dette punktet, at $\beta = 2$.

Kva er sannsynet for at ei tilfeldig vald snøprøve har densitet mellom 0,5 og 0,9 kg/dm^3?

b) Finn sannsynsmaksimeringsestimatoren for β basert på eit tilfeldig utval X_1, X_2, …, X_n av snødensitetar.

Kva blir estimatet dersom $n = 100$ og $\sum_{i=1}^{n} \ln(1 - x_i) = -104{,}0$?

Oppgåve 2 Temperatur i mars og april

I år (2012) var det mange stader i Noreg kaldare i april enn i mars.

La X vere gjennomsnittstemperaturen i mars og Y gjennomsnittstemperaturen i april ved Værnes eit tilfeldig valt år, begge målt i °C. Gå ut frå at X er normalfordelt med forventningsverdi μ_{m} og varians σ^2, og at Y er normalfordelt med forventningsverdi μ_{a} og varians σ^2.

a) Gå ut frå at vi har data i form av verdiar av X og Y for eit tilfeldig utval av år. Forklar korleis du grafisk kan undersøkje om føresetnaden om normalfordeling held. Korleis kan du grafisk undersøkje om X og Y er uavhengige?

Skisser òg grovt korleis slike grafiske framstillingar kan sjå ut både når føresetnaden om normalfordeling er oppfylt og når han ikkje er det, samt når X og Y er uavhengige og når dei ikkje er det.

Gjennomsnittstemperaturen i °C ved Værnes for åra 2001–2012 var slik:

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
x_i (mars)	$-2{,}5$	0,5	3,3	2,6	$-0{,}7$	$-4{,}6$	3,3	0,8	1,9	$-0{,}5$	1,2	3,8
y_i (april)	4,1	7,2	5,0	7,9	5,8	4,9	5,0	5,9	6,9	4,8	6,7	3,2

Det blir oppgitt at $\sum_{i=1}^{12} x_i = 9{,}10$, $\sum_{i=1}^{12} y_i = 67{,}40$, $\sum_{i=1}^{12} x_i^2 = 77{,}07$, $\sum_{i=1}^{12} y_i^2 = 399{,}30$, $\sum_{i=1}^{12}(x_i - y_i)^2 = 364{,}53$, der $i = 1$ står for år 2001, $i = 2$ for 2002 osv.

b) Gå ut frå at mars-temperaturane frå år til år er uavhengige. Finn eit 99 %-konfidensintervall for forventa gjennomsnittstemperatur i mars.

Vi ønskjer ved hypotesetesting å prøve å påvise at differansen mellom forventa gjennomsnittstemperatur i april og mars er mindre enn 5 °C. Gå ut frå at også april-temperaturane er uavhengige frå år til år.

c) Sett opp hypotesane.

Vi kan anten bruke ein toutvalstest eller ein partest. Kva vil du gjere? Argumenter for valet ditt, og utfør testen du vel. Bruk signifikansnivå $\alpha = 0{,}05$.

Oppgåve 3 Kjemisk fabrikk

Ei maskin på ein fabrikk utfører ein prosedyre for å lage eit kjemikalium. Eit giftig biprodukt blir danna i ei mengd på X gram kvar gong maskina utfører prosedyren, der X er normalfordelt med forventningsverdi 20 og standardavvik 4. Dersom meir enn 25 g av biproduktet blir danna, lyser ei varsellampe til prosedyren er avslutta. Maskina kan innstillast slik at ho utfører prosedyren fleire gonger på rad, men ho kan ikkje stoppast før alt er ferdig. Mengda biprodukt er uavhengig frå gong til gong.

a) Kva er sannsynet for at lampa lyser når maskina utfører prosedyren éin gong?

Maskina utfører prosedyren tre gonger. Kva er sannsynet for at lampa lyser minst éin gong?

Maskina utfører prosedyren 100 gonger. Finn eit tilnærma sannsyn for at lampa lyser 15 eller fleire gonger.

b) Forureiningsstyresmaktene krev at sannsynet for at maskina produserer 500 g eller meir av biproduktet på ein dag, skal vere 0,01 eller mindre. Kor mange gonger kan maskina utføre prosedyren på ein dag for at dette skal vere oppfylt?

Oppgåve 4 Grillfest i kveld?

Ein gjeng studentar vil ta seg ein velfortent kveld fri frå eksamensførebuingar etter statistikk-eksamen. Dei ønskjer å grille, men med vêratterhald: Det skal verken regne eller vere for kaldt (under 15 °C).

a) La A vere hendinga at det regnar, B hendinga at det er kaldt (under 15 °C) og C hendinga at det er grillvêr (15 °C eller meir og ikkje regn).

Framstill hendingane i eit Venn-diagram.

Vidare veit vi at $P(A) = P(B) = 0{,}4$, og $P(A \cap B) = 0{,}2$.

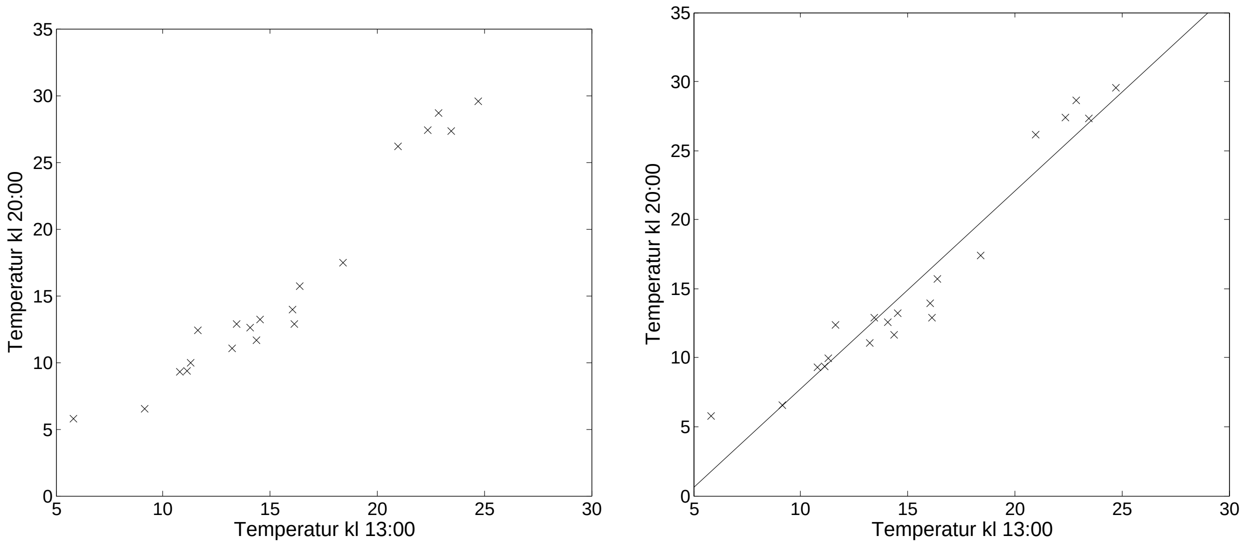
Er A og B disjunkte? Er A og B uavhengige?

Finn $P(C)$.

Er A og C disjunkte? Finn $P(C \mid A')$.

(Grunngje svara kort.)

Dei utset avgjersla til rett etter eksamen, klokka 13, og vil vurdere om det blir grillfest basert på vêret då. For å gjere eit best mogleg val, lagar dei ein enkel lineær regresjonsmodell med temperatur klokka 13, x, som forklaringsvariabel (uavhengig variabel) og temperatur klokka 20, Y, som responsvariabel (avhengig variabel), begge målt i °C: $Y = \alpha + \beta x + \epsilon$, der ϵ er normalfordelt med forventningsverdi 0 og varians σ^2.



Figur 1: Temperatur kl. 13 og kl. 20 på 26. mai i åra 1992–2011. Estimert regresjonslinje er teikna inn til høgre.

La x_i vere temperaturen klokka 13 og y_i temperaturen klokka 20 på 26. mai i åra 1992–2011, der $i = 1$ står for år 1992, $i = 2$ for 1993 osv. ($n = 20$ år). Hybelverten til ein av studentane har data tilgjengeleg (figur 1, venstre). Vi går ut frå at dataa er eit tilfeldig utval frå den nemnde regresjonsmodellen.

La $\hat{\alpha}$ og $\hat{\beta}$ vere estimatorane av α og β som ein får ved minste kvadrats metode. I punktet nedanfor kan du utan bevis bruke estimatorane oppgitt i *Tabeller og formler i statistikk*, og at $\hat{\beta}$ er forventningsrett, normalfordelt, har varians $\sigma^2 / \sum_{i=1}^n (x_i - \bar{x})^2$ og er uavhengig av $\bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$.

b) Finn sannsynsfordeling, forventningsverdi og varians til $\hat{\alpha}$. Grunngje kort svara og overgangane i utrekningane.

List opp føresetnadene som er gjorde i regresjonsmodellen. Vurder om mogleg frå figur 1 om desse er oppfylte.

Estimata baserte på dataa blir $\hat{\alpha} = -6{,}57$, $\hat{\beta} = 1{,}43$ og $\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}x_i)^2 = 2{,}04^2$. Vidare er $\bar{x} = 15{,}5$ og $\sum_{i=1}^n (x_i - \bar{x})^2 = 510{,}7$. Den estimerte regresjonslinja er teikna inn i figur 1 (høgre).

c) Finn eit estimat av forventa temperatur klokka 20 dersom temperaturen klokka 13 er 15 °C.

Finn eit 95 %-prediksjonsintervall for temperaturen klokka 20 dersom temperaturen klokka 13 er 15 °C.