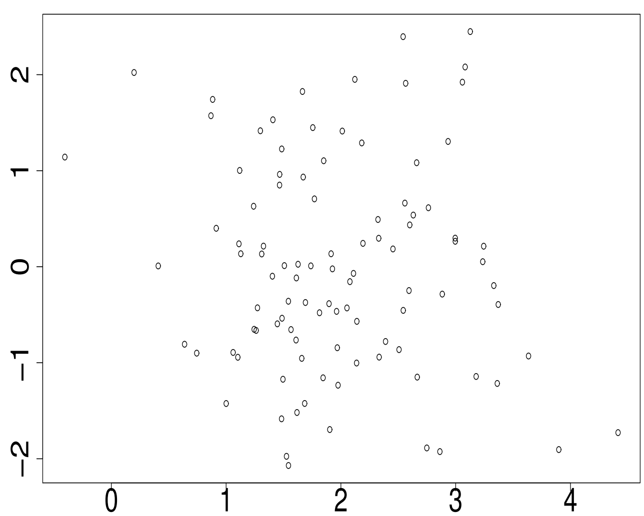
TMA4245 Statistikk
Eksamen juni 2015

Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Oppgave 1



Figur 1: Parvise data frå ei simultanfordeling.

Figur 1 viser parvise utfall av to stokastiske variablar X (horisontalt) og Y (vertikalt).

a) Gje eit anslag på korrelasjonen mellom X og Y. Skriv ei kort grunngjeving.

Forventingsverdien og standardavviket for kvar av dei to variablane har heiltalsverdiar. Gje eit anslag for desse verdiane ved å studere Figur 1 visuelt.

Oppgave 2

Gå ut i frå at A og C er uavhengige hendingar, og at B og C er uavhengige hendingar.

a) Kan $A \cap B$ og C vere avhengige? Kan $A \cup B$ og C vere avhengige? (Vink: Sjå til dømes på tilfellet $A \cap B \cap C = \emptyset$.)

Kan $A \cup B$ og C være avhengige dersom A og B er disjunkte?

Oppgave 3

Ein farmasøytisk fabrikk produserer ein medisin i flytande form som skal seljast på flasker med oppgjeve innhald. Denne medisinen inneheld ein viktig komponent som krev at ein fÃ¸retek ein jamleg kontroll av kvaliteten. Dette blir gjort ved at ein for kvar produksjonsserie tek ut eit tilfeldig utval flasker der innhaldet blir analysert. Kvar gong ein serie prøver blir analysert, analyserer ein samstundes ei kontrolløysing med kjent konsentrasjon 0.10 mg/l. Dette gjer ein for å sikre at analysemetoden er rett kalibrert. Sidan analysemetoden ikkje er heilt presis, vil den målte konsentrasjonen variere. Utfallet av analysen av kontrolløysinga kan reknast som ein normalfordelt stokastisk variabel X med forventingsverdi μ og varians σ^2, der σ^2 er variansen i målefeilen for analysemetoden og μ under normale omstende er lik 0.10 mg/l.

Ei *alarmhending* A for analysemetoden er definert ved at den målte verdien X i kontrolløysinga avvik meir enn eit standardavvik frå konsentrasjonen 0.10 mg/l, altså $|X - 0.1| > \sigma$ (dvs. $X < 0.1 - \sigma$ eller $X > 0.1 + \sigma$). Ei *aksjonshending* B er definert ved at den målte verdien X i kontrolløysinga avvik meir enn 2 standardavvik frå 0.10 mg/l, dvs. $|X - 0.1| > 2\sigma$.

a) Gå ut i frå (berre i dette punktet) at $\sigma = 0.01$ mg/l.

Rekn ut $P(B)$ og $P(B \mid A)$ når $\mu = 0.10$ mg/l.

Gå ut i frå at ei forureining har kome inn i prøva, slik at $\mu = 0.11$ mg/l, og rekn ut $P(B)$.

Gå i resten av oppgåva ut i frå at $\mu = 0.10$ mg/l. For å estimere σ^2 skal vi nytte måleresultata $x_1, x_2, \ldots, x_n$, frå n uavhengige analyser av kontrolløysinga. Vi kan derfor sjå på $x_1, x_2, \ldots, x_n$ som utfall frå n uavhengige stokastiske variablar $X_1, X_2, \ldots, X_n$, der kvar X_i, $i = 1, \ldots, n$, har same fordeling som den stokastiske variabelen X.

b) Vis at sannsynsmaksimeringsestimatoren (SME) for σ^2 er

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

SjÃ¸lv om μ er kjent, kan ein og bruke estimatoren

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

der $\overline{X} = \sum_{i=1}^{n} X_i/n$.

I dei neste punkta kan du bruke utan bevis at

$$\sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2 \text{ og } \sum_{i=1}^{n}(X_i - \overline{X})^2/\sigma^2$$

er χ^2-fordelte(kjikvadratfordelte) med n og $n-1$ fridomsgrader i gitt rekkjefÃ¸lge.

c) Vis at både $\widehat{\sigma}^2$ og S^2 er forventingsrette.

Kven av dei to estimatorane vil du velje? Grunngjev svaret.

Resultata målt i mg/l av 20 analyser av kontrolløysinga har gjeve $\sum_{i=1}^{20} x_i = 1.9240$ og $\sum_{i=1}^{20} x_i^2 = 0.1866$.

d) Bestem eit 90 %-konfidensintervall for σ^2 ved å nytte estimatoren du kom fram til var den beste i (c).

Oppgave 4

Let $X_1, ..., X_n$ vere eit tilfeldig utval (uavhengige, identisk fordelte stokastiske variablar) frå ei eksponensialfordeling med ukjent parameter μ, dvs. frå ei fordeling med sannsynstettleik

$$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}\ ;\ x \geq 0,\ \mu > 0.$$

Nullhypotesen $H_0 : \mu = 1$ skal testast mot alternativet $H_1 : \mu < 1$. I fÃ¸rste omgang skal du nytte testobservatoren

$$T = \min\{X_1, ..., X_n\},$$

a) Vis at sannsynstettleiken til T er

$$f_T(t) = \frac{n}{\mu} e^{-\frac{nt}{\mu}}\ ;\ t \geq 0.$$

Bestem forventningsverdien og variansen til T.

Testen som baserer seg på bruk av T, forkastar H_0 for små verdiar av T. Dette kan ein formulera på denne måten:

Test1: Dersom $T < c_1$, forkastar ein H_0.

Signifikansnivået for testane i denne oppgåva er α.

b) Vis at

$$c_1 = \frac{1}{n} \ln \frac{1}{1-\alpha}.$$

SME for μ er utvalsgjennomsnittet $\overline{X} = \Sigma_{i=1}^n X_i/n$.

c) Formuler ein Test2 basert på $\overline{X}$. Bestem ei tilnærma sannsynsfordeling til $\overline{X}$, og finn (tilnærma) det kritiske området for Test2 dersom du kan gå ut i frå at $n \geq 30$, .

d) Finn eit uttrykk for teststyrken for kvar av dei to testane ($n \geq 30$). (Teststyrke er sannsynet for å forkaste H_0 når H_0 ikkje er rett; styrken avheng av eit skjerpa alternativ.)

Kva for ein test er best? Kvifor? Du får lov til å basere konklusjonen din på resultat rekna ut for $\alpha = 0.05$, $n = 30$, og skjerpa alternativ hypotese $H_1' : \mu = 0.8$.

Kvifor er Test1 ein dårleg test?

Fasit

1. a) 0,2,0,1,1

2. a) Ja, Ja, Nei

3. a) 0.0456, 0.1437, 0.16 **c)** Foretrekker $\widehat{\sigma}^2$ **d)** $[0.000057, 0.000166]$

4. a) $\mathrm{E}[T] = \mu/n$, $\mathrm{Var}[T] = \mu^2/n^2$ **c)** Forkast H_0 dersom $\overline{X} \leq 1 - \frac{z_\alpha}{n^{1/2}}$.