Oppgave

Anta Y_1 og Y_2 uavhengige, $Y_1 \sim \mathcal{N}(\mu_1 \sigma)$, og $Y_2 \sim \mathcal{N}(\mu_2, \sigma)$. Definer

$$Z_1 = \frac{Y_1 + Y_2}{2}, \qquad Z_2 = \frac{Y_1 - Y_2}{2}.$$

Vi at Z_1 og Z_2 er uavhengige.

Fremgangsmåte

Det er snakk om å vise at kontinuerlige variable er uavhengige, så vi ønsker å vise at

$$g_{Z_1,Z_2}(z_1, z_2) = g_{Z_1}(z_1) g_{Z_2}(z_2).$$

Her har konstanter ikke noe å si, fordi alle tre funksjonene må integrere til 1, så eventuelle konstanter fikser vi til slutt. Vi bruker tegnet $\propto$ for å si at ting er proposjonalt med hverandre: "Det samme som, bortsett fra at vi må gange med en konstant".

For å løse oppgaven finner vi først $g_{Z_1,Z_2}(z_1, z_2)$, deretter skriver vi om uttrykket til vi har høyresiden av likningen over. Da vil vi ha funnet både $g_{Z_1}(z_1), g_{Z_2}(z_2)$, og at Z_1 og Z_2 er uavhengige.

Merk: Z_1, Z_2 er ikke standard normalfordelte. Her brukes bare Z som et variabelnavn. Vi vet ikke forventningsverdien eller variansen (men vi kunne greit regnet det ut).

Løsning

Finn $g_{Z_1,Z_2}(z_1, z_2)$

Vi skal nå transformere variablene fra Y_i til Z_i. Hvis vi har et formelhefte kan vi slå opp i det. La oss nå prøve å finne ut av det uten å slå det opp.

$$\begin{aligned} P(Z_1 < z_1, Z_2 < z_2) &= P(Y_1 < y_1, Y_2 < y_2) \\ G_{Z_1,Z_2}(z_1, z_2) &= F_{Y_1,Y_2}(y_1, y_2) \\ g_{Z_1,Z_2}(z_1, z_2) &= f_{Y_1,Y_2}(y_1, y_2)\ \frac{dy}{dz} \end{aligned}$$

Kommentar: Med $\frac{dy}{dz}$ mener vi her (den positive) Jakobi-determinanten, som ofte skrives $|J|$.

Før vi kan finne den deriverte må vi skrive opp funksjonene:

$$\begin{aligned} Y_1 &= Z_1 + Z_2 \\ Y_2 &= Z_1 - Z_2 \end{aligned}$$

og så:

$$\begin{aligned} \frac{dy}{dz} &= \left\| \begin{bmatrix} \frac{dy_1}{dz_1} & \frac{dy_1}{dz_2} \\ \frac{dy_2}{dz_1} & \frac{dy_2}{dz_2} \end{bmatrix} \right\| \\ &= \left\| \begin{bmatrix} 1 & 1 \\ 1 & \text{-1} \end{bmatrix} \right\| = |-1-1| = 2. \end{aligned}$$

Da får vi

$$\begin{aligned} g_{Z_1,Z_2}(z_1, z_2) &= f_{Y_1,Y_2}(y_1, y_2)\ \frac{dy}{dz} \\ &= \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(y_1-\mu_1)^2}{2\sigma^2}\right] \cdot \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(y_2-\mu_2)^2}{2\sigma^2}\right] \cdot 2 \\ &= \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(z_1+z_2-\mu_1)^2}{2\sigma^2}\right] \cdot \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(z_1-z_2-\mu_2)^2}{2\sigma^2}\right] \cdot 2 \end{aligned}$$

Dette kunne vært en oppgave i seg selv, men vi er ikke ferdige.

Skriv $g_{Z_1,Z_2}(z_1, z_2)$ som et produkt av to funksjoner

$$
\begin{aligned}
g_{Z_1,Z_2}(z_1, z_2) = f_{Y_1,Y_2}(y_1, y_2)\ \frac{dy}{dz} \\
&\propto exp\left[-\frac{(z_1+z_2-\mu_1)^2}{2\sigma^2}\right]\cdot exp\left[-\frac{(z_1-z_2-\mu_2)^2}{2\sigma^2}\right] \\
&\propto exp\left[-\frac{1}{2\sigma^2}\left[(z_1+z_2-\mu_1)^2+(z_1-z_2-\mu_2)^2\right]\right] \\
&\propto exp\left[-\frac{1}{2\sigma^2}\left[2z_1^2-2z_1\mu_1-2z_1\mu_2\right]\right] \\
&\qquad\cdot exp\left[-\frac{1}{2\sigma^2}\left[2z_2^2-2z_2\mu_1+2z_2\mu_2\right]\right] \\
&\qquad\cdot exp\left[-\frac{1}{2\sigma^2}\left[-\mu_1^2+\mu_2^2\right]\right] \\
&\propto exp\left[-\frac{1}{2\sigma^2}\left[2z_1^2-2z_1\mu_1-2z_1\mu_2\right]\right] \\
&\qquad\cdot exp\left[-\frac{1}{2\sigma^2}\left[2z_2^2-2z_2\mu_1+2z_2\mu_2\right]\right] \\
&= h(z_1)\cdot q(z_2),
\end{aligned}
$$

hvor h og q er definert slik at det passer; slik at q integrerer til 1, og at de andre konstantene er tatt inn i h. Merk at μ'ene er kjente konstanter, og derfor ikke med i funksjonsuttrykkene.

Avsluttende argumenter

Vi vet at

$$
\begin{aligned}
g_{Z_1}(z_1) &= \int g_{Z_1,Z_2}(z_1, z_2)\ \mathrm{d}z_2 \\
&= \int h(z_1)\cdot q(z_2)\ \mathrm{d}z_2 \\
&= h_1(z_1)\cdot 1,
\end{aligned}
$$

altså er h marginalfordelingen til Z_1. Tilsvarende kan vi vise at q er marginalfordelingen til Z_2. Da har vi vist det vi skulle.

Hvis vi hadde prøvd hardt kunne vi funnet uttrykket til h og q, men det gidder vi ikke, for vi har allerede svart på oppgaven.

Vi kunne argumentert litt annerledes på slutten. Det viktige er at den 2-variable funksjonen $g_{Z_1,Z_2}(z_1, z_2)$ kan skrives som et produkt av to 1-variable funksjoner, $h(z_1)$ og $q(z_2)$.