Extra uitleg bij sectie 5.4

Veronderstel dat X en Y twee continue stochasten zijn met simultane verdeling $f(x, y)$, waarvoor al dan niet een expliciete uitdrukking bestaat. Kunnen we de dichtheid van $g(X, Y)$, met g een functie, bepalen? De uitleg in het boek is niet wiskundig verantwoord, dus zullen we in dit document een alternatieve (en sterk aanbevolen!) methode geven waarmee we dit probleem kunnen aanpakken. De methode beschreven in dit document is ook op het hoorcollege besproken.

De dichtheid van $g(X, Y)$ is te bepalen met behulp van de volgende stappen:

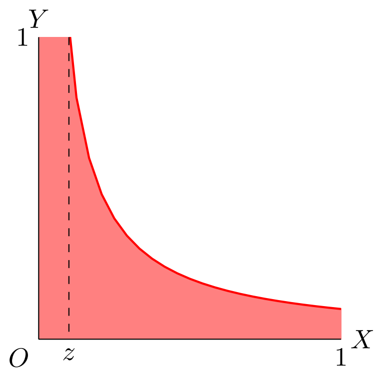
Stappenplan voor het bepalen van de dichtheid van $g(X, Y)$ als de simultane dichtheid van X en Y een expliciete uitdrukking heeft

1. Schets de verzameling $g(X, Y) \le z$ in het (X, Y)-vlak. Deze verzameling noemen we R.
2. Bereken de kans $P\left(g(X, Y) \le z\right) = \iint_R f(x, y)\, dA$. Gebruik hiervoor de methoden die je bij Infi B hebt geleerd.
3. De dichtheid van $g(X, Y)$ is $\frac{d}{dz} P\left(g(X, Y) \le z\right)$.

We laten zien hoe dit werkt aan de hand van een aantal voorbeelden.

Voorbeeld. Stel dat X en Y onafhankelijk en uniform verdeeld zijn op $(0, 1)$. We willen de dichtheid van $Z = XY$ bepalen.

1. We laten $0 < z < 1$, omdat het duidelijk is dat $0 < Z < 1$ geldt. De verzameling $XY \le z$ ziet er dan zo uit:



2. De marginale verdeling van X en Y is 1 op $(0, 1)$ en 0 daarbuiten. Omdat X en Y onafhankelijk zijn, is de simultane verdeling van X en Y 1 op $(0, 1) \times (0, 1)$ en 0 daarbuiten. Deze integreren we over het gebied dat we hierboven geschetst hebben. Het is duidelijk dat we de integraal moeten opsplitsen in twee stukken en wel als volgt:

$$\begin{aligned} P(XY \le z) &= \iint_R dA \\ &= \int_0^z \int_0^1 dy dx + \int_z^1 \int_0^{z/x} dy dx \\ &= \int_0^z dx + \int_z^1 \frac{z}{x}\, dx \;=\; z - z \log z. \end{aligned}$$

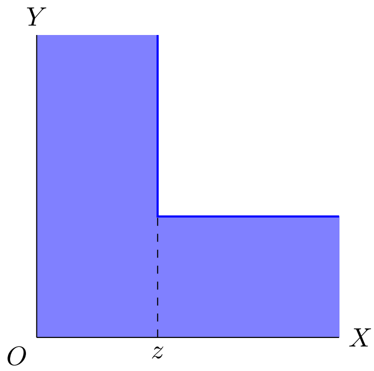
3. De dichtheid op $0 < z < 1$ vinden we door te differentiëren.
Aangezien $\frac{d}{dz}\left(z - z \log z\right) = -\log z$, volgt dat:

$$f_{XY}(z) = \begin{cases} -\log z, & 0 < z < 1; \\ 0, & \text{anders.} \end{cases}$$

□

Voorbeeld. Stel dat X en Y onafhankelijk en exponentieel verdeeld zijn met parameter λ. We willen de dichtheid van $Z = \min(X, Y)$ bepalen.

1. Nu geldt $Z \geq 0$, dus laten we $z \geq 0$. De verzameling $\min(X, Y) \leq z$ ziet er als volgt uit:



2. De simultane verdeling is $\lambda^2 e^{-\lambda(x+y)}$, welke we over bovenstaand gebied moeten integreren. Wederom splitsen we de integraal op:

$$\begin{aligned} P(\min(X,Y) \leq z) &= \int_0^z \int_0^\infty \lambda^2 e^{-\lambda(x+y)}\, dydx + \int_z^\infty \int_0^z \lambda^2 e^{-\lambda(x+y)}\, dydx \\ &= \int_0^z \lambda e^{-\lambda x}\, dx + \int_z^\infty -\lambda \left(e^{-\lambda(x+z)} - e^{-\lambda x}\right) dx \\ &= 1 - e^{-\lambda z} - e^{-2\lambda z} + e^{-\lambda z} \; = \; 1 - e^{-2\lambda z}. \end{aligned}$$

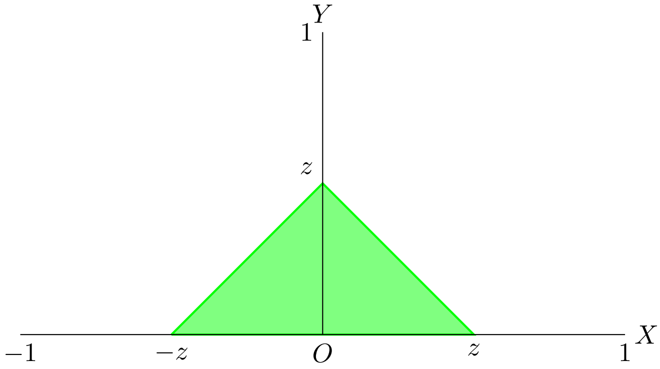
3. Door te differentiëren naar z vinden we de verdeling van $\min(X, Y)$:

$$f_{\min(X,Y)}(z) = \begin{cases} 2\lambda e^{-2\lambda z}, & z \geq 0; \\ 0, & \text{anders.} \end{cases}$$

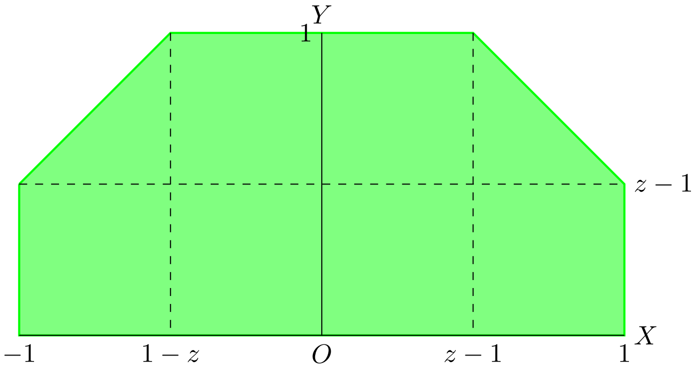
   We zien dat $\min(X, Y)$ exponentieel verdeeld is met parameter 2λ (vergelijk met opgave 5.2.9(c)).

Voorbeeld. Stel dat X uniform verdeeld is op $(-1,1)$, Y uniform verdeeld op $(0,1)$ en dat X en Y onafhankelijk zijn. We willen de dichtheid van $Z = |X| + |Y|$ bepalen.

1. Nu geldt $0 \leq Z < 2$, dus kiezen we $0 \leq z < 2$. We moeten twee gevallen onderscheiden: $z \leq 1$ en $z > 1$. Indien $z \leq 1$ ziet de verzameling $|X| + |Y| \leq 1$ er als volgt uit:



Voor $z > 1$ ziet de verzameling er echter zo uit:



2. We beginnen eerst met het geval $z \leq 1$. In dit geval geldt:

$$
\begin{aligned}
P(|X| + |Y| \leq z) &= \int_0^z \int_{y-z}^{z-y} \frac{1}{2}\, dx dy \\
&= \int_0^z z - y\, dy \;=\; \frac{z^2}{2}.
\end{aligned}
$$

Voor $z > 1$ geldt:

$$
\begin{aligned}
P(|X| + |Y| \leq z) &= \int_0^{z-1} \int_{-1}^{1} \frac{1}{2}\, dx dy + \int_{z-1}^{1} \int_{y-z}^{z-y} \frac{1}{2}\, dx dy \\
&= \int_0^{z-1} dy + \int_{z-1}^{1} z - y\, dy \\
&= z - 1 - \frac{z(z-2)}{2} \;=\; -\frac{z^2}{2} + 2z - 1.
\end{aligned}
$$

3. Differentiëren naar z geeft de dichtheid van $|X| + |Y|$:

$$
f_{|X|+|Y|}(z) = \begin{cases} z, & 0 \leq z \leq 1; \\ 2 - z, & 1 < z < 2; \\ 0, & \text{anders.} \end{cases}
$$

Zelfs als de simultane verdeling van X en Y niet expliciet gegeven is, kunnen we deze stappen gebruiken om iets te zeggen over de verdeling van $g(X, Y)$ in termen van de verdeling $f(x, y)$. Daartoe hebben we de volgende twee stellingen nodig. De eerste stelling formuleren we op een zeer niet-formele manier:

Stelling (Verwisselen van integraal en differentiaal). Bij het vak kansrekening geldt

$$\frac{d}{dz} \int_R f\left(g(x,z),\, h(x,z)\right)\, dx \quad = \quad \int_R \frac{\partial}{\partial z} f\left(g(x,z)\, h(x,z)\right)\, dx,$$

mits de rand van R niet afhangt van z.

De exacte formulering van deze stelling stel voorwaarden aan de functie f die op dit moment bij jullie nog niet bekend zal zijn (hiervoor is maattheorie nodig) en we daarom niet zullen vermelden. De tweede stelling is een stelling die bij Infi A aan bod is geweest (hoewel misschien niet in deze vorm):

Stelling (Hoofdstelling der infinitesimaalrekening). Zij $f : \mathbb{R} \to \mathbb{R}$ continu en $g, h : \mathbb{R} \to \mathbb{R}$ differentieerbaar, dan geldt:

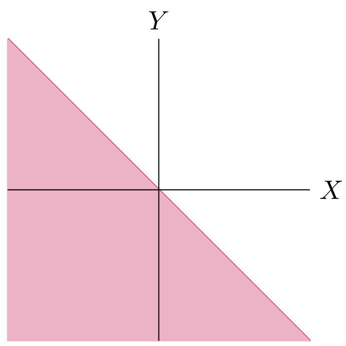
$$\frac{d}{dt} \int_{g(t)}^{h(t)} f(x)\, dx \quad = \quad f\left(h(t)\right) h'(t) - f\left(g(t)\right) g'(t).$$

De stelling kan uiteraard gegeneraliseerd worden, maar bij dit vak is de formulering zoals hierboven voldoende.

De volgende twee voorbeelden laten zien hoe we deze stellingen kunnen gebruiken.

Voorbeeld (Convolutieformule). Zij X en Y continue stochasten met simultane verdeling $f(x, y)$. Bepaal de verdeling van $Z = X + Y$.

1. We nemen aan dat X en Y op heel $\mathbb{R}$ bestaan. Dit geldt dan uiteraard ook voor Z en de verzameling $X + Y \leq z$ voor zekere $z \in \mathbb{R}$ ziet er als volgt uit:



2. De kans is dus $P(X + Y \leq z) = \displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x, y)\, dydx.$

3. De dichtheid vinden we door dit laatste te differentiëren, dus:

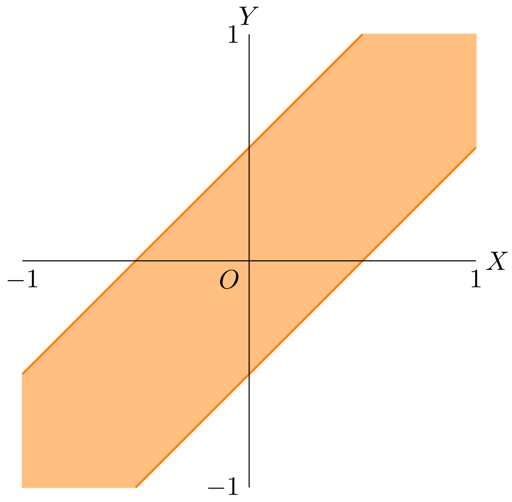
$$\begin{aligned}
f_{X+Y}(z) &= \frac{d}{dz} \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x,y)\, dydx \\
&= \int_{-\infty}^{\infty} \frac{\partial}{\partial z} \int_{-\infty}^{z-x} f(x,y)\, dydx && \text{vanwege de eerste stelling} \\
&= \int_{-\infty}^{\infty} f(x, z-x) \cdot 1\, dx && \text{vanwege de tweede stelling,} \\
& && \text{gebruik ook dat als } x, y \to \pm\infty \text{, dan } f(x,y) \to 0 \\
& && \text{voor alle kansdichtheden}^{1} \\
&= \int_{-\infty}^{\infty} f(x, z-x)\, dx.
\end{aligned}$$

Hiermee is de convolutieformule bewezen.

[1]Dit geldt omdat $f(x, y)$ niet-negatief moet zijn en bovendien $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dxdy = 1$ moet gelden.

Voorbeeld. Zij X en Y continue stochasten met simultane verdeling $f(x,y)$. Bepaal de verdeling van $Z = |X-Y|$.

1. De stochast Z is niet-negatief. De verzameling $|X-Y| \leq z$ er als volgt uit:



2. De cumulatieve kans is

$$P\left(|X-Y| \leq z\right) \quad = \quad \int_{-\infty}^{\infty} \int_{x-z}^{x+z} f(x,y)\, dy dx.$$

3. De dichtheid van $|X-Y|$ vinden we door te differentiëren:

$$\begin{aligned} \frac{d}{dz} \int_{-\infty}^{\infty} \int_{x-z}^{x+z} f(x,y)\, dy dx &= \int_{-\infty}^{\infty} \frac{\partial}{\partial z} \int_{x-z}^{x+z} f(x,y)\, dy dx && \text{vanwege de eerste stelling} \\ &= \int_{-\infty}^{\infty} f(x,x+z) \cdot 1 - f(x,x-z) \cdot -1\, dx && \text{vanwege de tweede stelling} \\ &= \int_{-\infty}^{\infty} f(x,x+z) + f(x,x-z)\, dx. \end{aligned}$$

De dichtheid is dus:

$$f_{|X-Y|}(z) \quad = \quad \begin{cases} \displaystyle\int_{-\infty}^{\infty} f(x,x+z) + f(x,x-z)\, dx, & z \geq 0; \\ 0, & \text{anders.} \end{cases}$$