Sums of independent random variables

Q. X, Y independent random variables. What is the distribution of $X+Y$?

Let's find the cdf when X and Y are (jointly) continuous:

$$F_{X+Y}(a) = \mathbb{P}(X+Y \le a) = \iint_{x+y \le a} f_{X,Y}(x,y)\, dx\, dy$$

$$= \iint_{x+y \le a} f_X(x) f_Y(y)\, dx\, dy \qquad \text{independence}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)\, dx \; f_Y(y)\, dy$$

$$= \int_{-\infty}^{\infty} F_X(a-y) f_Y(y)\, dy$$

Differentiate (in a):

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y)\, dy \qquad \text{"convolution of } f_X \text{ and } f_Y\text{"}$$

Ex. $X, Y \sim U(0,1)$ independent.

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y)\, dy$$

$$= \int_0^1 f_X(a-y)\, dy \quad ; \quad f_Y(y) = \begin{cases} 0 & \text{unless} \\ 1 & 0 < y < 1 \end{cases}$$

$$= 0 \text{ unless} \quad ; \quad f_X(a-y) = \begin{cases} 0 & \text{unless} \\ 1 & 0 < a-y < 1 \end{cases}$$

$$\Updownarrow$$

$$a-1 < y < a$$

If $a<0$, $\quad a-1 \le y \le a$ implies $y<0$. So

$$\int_0^1 f_X(a-y)\,dy = 0$$

If $0 \le a \le 1$, $\quad \left.\begin{array}{l} a-1 \le y \le a \\ 0 \le y \le 1 \end{array}\right] \iff 0 \le y \le a$. In this case,

$$\int_0^1 f_X(a-y)\,dy = \int_0^a 1\,dy = a$$

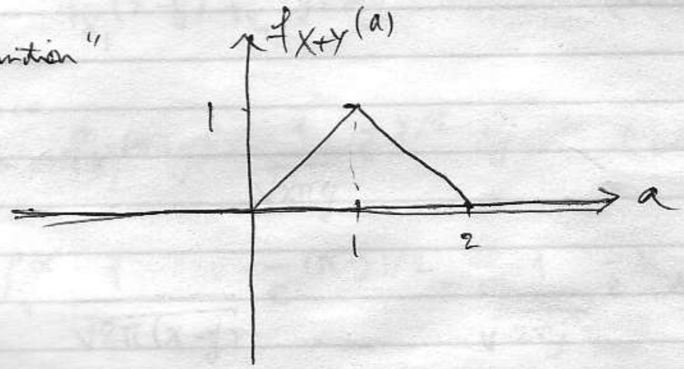
If $1<a<2$, $\quad \left.\begin{array}{l} a-1 \le y \le a \\ 0 \le y \le 1 \end{array}\right] \iff a-1 \le y \le 1$. Here,

$$\int_0^1 f_X(a-y)\,dy = \int_{a-1}^1 dy = 1-(a-1) = 2-a$$

If $a>2$, $\quad \int_0^1 f_X(a-y)\,dy = 0$.

$$\therefore \quad f_{X+Y}(a) = \begin{cases} a, & 0 \le a \le 1 \\ 2-a, & 1 \le a \le 2 \\ 0, & \text{otherwise} \end{cases}$$

"triangular density function"



Example. X_1, X_2 independent $N(0,1)$'s.

(a) Find $f_{X_1^2}$.

(b) Find $f_{X_1^2+X_2^2}$.

Sol^n to (a) $\quad \mathbb{P}(X_1^2 \le a) = \mathbb{P}(-\sqrt{a} \le X_1 \le \sqrt{a}) \qquad \underline{\underline{a \ge 0}}$

$$= \Phi(\sqrt{a}) - \Phi(-\sqrt{a})$$

$$= 2\Phi(\sqrt{a}) - 1$$

$$\therefore \quad f_{X_1^2}(a) = 2 \cdot \frac{1}{2\sqrt{a}}\, \Phi'(\sqrt{a})$$

$$= \frac{1}{\sqrt{2\pi a}}\, e^{-a/2} \qquad a > 0$$

If $a < 0$, $f_{X_1^2}(a) = 0$, since $\mathbb{P}(X_1^2 \le a) = 0$ then.

Sol^n to (b) $\quad f_{X_1^2+X_2^2}(x) = \int_{-\infty}^{\infty} f_{X_1}(x-y)\, f_{X_2}(y)\, dy \qquad$ (convolution)

$$= \int_0^{\infty} f_{X_1}(x-y)\, \frac{1}{\sqrt{2\pi y}}\, e^{-y/2}\, dy \qquad \text{(part (a))}$$

$$= \int_0^{x} \frac{1}{\sqrt{2\pi(x-y)}}\, e^{-(x-y)/2} \cdot \frac{1}{\sqrt{2\pi y}}\, e^{-y/2}\, dy \qquad \text{(part (a))}$$

$$= \frac{e^{-x/2}}{2\pi} \int_0^{x} \frac{1}{\sqrt{y(x-y)}}\, dy$$

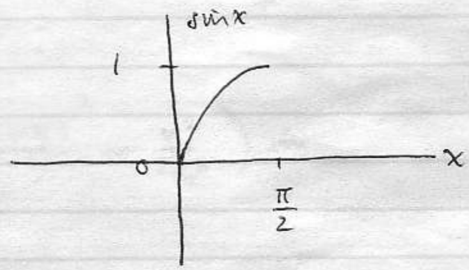
$\sqrt{y(x-y)} = x\sqrt{\frac{y}{x}\left(1-\frac{y}{x}\right)}$, since $x>0$. Therefore,

$$f_{X_1^2+X_2^2}(x) = \frac{e^{-x/2}}{2\pi x}\int_0^x \frac{dy}{\sqrt{\frac{y}{x}\left(1-\frac{y}{x}\right)}}$$

$$= \frac{e^{-x/2}}{2\pi}\int_0^1 \frac{dz}{\sqrt{z(1-z)}} \qquad (z = y/x)$$

$$= \frac{e^{-x/2}}{\pi}\int_0^1 \frac{dw}{\sqrt{1-w^2}} \qquad (w = \sqrt{z})$$

$$= \frac{e^{-x/2}}{\pi}\left(\sin^{-1}1 - \sin^{-1}0\right)$$



$$= \frac{e^{-x/2}}{\pi}\left(\frac{\pi}{2} - 0\right)$$

$$= \frac{1}{2}e^{-x/2}$$

$$\Rightarrow \quad f_{X_1^2+X_2^2}(x) = \begin{cases} \frac{1}{2}e^{-x/2} & , x>0 \\ 0 & , x\le 0 \end{cases}$$

i.e. $\boxed{X_1^2+X_2^2 \sim \mathcal{E}xp\left(\tfrac{1}{2}\right)}$.

Sums of independent normals.

1// X_1, X_2 indep't. $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$, resp.
Find the pdf of $X_1 + X_2$.

$$f_{X_1+X_2}(a) = \int_{-\infty}^{\infty} f_{X_1}(a-u)\, f_{X_2}(u)\, du$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma_1} \exp\left(-\frac{(a-u-\mu_1)^2}{2\sigma_1^2}\right) \frac{1}{\sqrt{2\pi}\,\sigma_2} \exp\left(-\frac{(u-\mu_2)^2}{2\sigma_2^2}\right) du$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(a-u-\mu_1)^2}{2\sigma_1^2} - \frac{(u-\mu_2)^2}{2\sigma_2^2}\right) du$$

$v = u - \mu_2$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(a-(\mu_1+\mu_2)-v)^2}{2\sigma_1^2} - \frac{v^2}{2\sigma_2^2}\right) dv$$

Let $\mu = \mu_1 + \mu_2$, for simplicity.

$$f_{X_1+X_2}(a) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(a-\mu-v)^2}{2\sigma_1^2} - \frac{v^2}{2\sigma_2^2}\right) dv$$

$$\frac{(a-\mu-v)^2}{2\sigma_1^2} + \frac{v^2}{2\sigma_2^2} = \frac{(a-\mu)^2 + v^2 - 2(a-\mu)v}{2\sigma_1^2} + \frac{v^2}{2\sigma_2^2}$$

$$= \frac{v^2}{2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right) + \frac{(a-\mu)^2}{2\sigma_1^2} - \left(\frac{a-\mu}{\sigma_1^2}\right)v$$

Let $\sigma^2 = \sigma_1^2 + \sigma_2^2$, for simplicity.

$$\frac{(a-\mu-v)^2}{2\sigma_1^2} + \frac{v^2}{2\sigma_2^2} = \frac{v^2}{2}\cdot\frac{\sigma^2}{\sigma_1^2\sigma_2^2} - 2\left(\frac{a-\mu}{2\sigma_1^2}\right)v + \frac{(a-\mu)^2}{2\sigma_1^2}$$

$$= \frac{1}{2\sigma_1^2}\left[v^2\cdot\frac{\sigma^2}{\sigma_2^2} - 2(a-\mu)v\right] + \frac{(a-\mu)^2}{2\sigma_1^2}$$

$$= \frac{1}{2\sigma_1^2}\left[v^2 \cdot \frac{\sigma^2}{\sigma_2^2} - 2\frac{(a-\mu)\sigma_2}{\sigma}\left(v\frac{\sigma}{\sigma_2}\right)\right] + \frac{(a-\mu)^2}{2\sigma_1^2}$$

$$= \frac{1}{2\sigma_1^2}\left[\left(\frac{v\sigma}{\sigma_2}\right)^2 - 2\frac{(a-\mu)\sigma_2}{\sigma}\left(v\frac{\sigma}{\sigma_2}\right) + (a-\mu)^2\frac{\sigma_2^2}{\sigma^2}\right] - \frac{(a-\mu)^2\sigma_2^2}{2\sigma^2\sigma_1^2} + \frac{(a-\mu)^2}{2\sigma_1^2}$$

$$= \frac{1}{2\sigma_1^2}\left(\frac{v\sigma}{\sigma_2} - (a-\mu)\frac{\sigma_2}{\sigma}\right)^2 + \frac{(a-\mu)^2}{2\sigma_1^2}\left[1 - \frac{\sigma_2^2}{\sigma^2}\right]$$

$$= \frac{1}{2\sigma_1^2}\left[\frac{v\sigma}{\sigma_2} - \frac{(a-\mu)\sigma_2}{\sigma}\right]^2 + \frac{(a-\mu)^2}{2\sigma^2} \qquad \left[\sigma^2 \triangleq \sigma_1^2 + \sigma_2^2\right]$$

$\therefore$

$$f_{X_1+X_2}(a) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{1}{2\sigma_1^2}\left(\frac{v\sigma}{\sigma_2} - \frac{(a-\mu)\sigma_2}{\sigma}\right)^2\right] dv \; . \; e^{-\frac{(a-\mu)^2}{2\sigma^2}}$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2} e^{-w^2/2\sigma_1^2}\, dw \cdot \frac{\sigma_2}{\sigma} \cdot \exp\left[-\frac{(a-\mu)^2}{2\sigma^2}\right] \qquad \left[w = \frac{v\sigma}{\sigma_2} - (\dots)\right]$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{(a-\mu)^2}{2\sigma^2}\right].$$

$\Rightarrow$ $X_i \sim N(\mu_i, \sigma_i^2)$ indep't. $\Rightarrow$ $X_1 + X_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

$\Rightarrow$ $\underbrace{X_1 + X_2}_{N(\mu_1+\mu_2,\, \sigma_1^2+\sigma_2^2+\sigma_3^2)} + X_3 \sim N(\mu_1+\mu_2+\mu_3, \sigma_1^2+\sigma_2^2+\sigma_3^2)$

<u>Theorem.</u> $X_i \sim N(\mu_i, \sigma_i^2)$, $i \le n$, independent.

Then $\sum_{j=1}^{n} X_j \sim N\left(\sum_1^n \mu_i, \sum_1^n \sigma_i^2\right)$.

<u>Summing Independent r.v.'s in the discrete case.</u>

X_1, X_2 independent. $\quad \mathbb{P}(X=x, Y=y) = p(x,y) = \mathbb{P}(X=x)\,\mathbb{P}(Y=y) = p_X(x)\,p_Y(y)$.

$$\left.\begin{aligned} p_{X_1+X_2}(k) = \mathbb{P}(X_1+X_2=k) &= \mathbb{P}(X_1=0, X_2=k) + \mathbb{P}(X_1=1, X_2=k-1) \ldots \\ &= p_{X_1}(0)\,p_{X_2}(k) + p_{X_1}(1)\,p_{X_2}(k-1) + \ldots \end{aligned}\right\}\text{ pos. integer case}$$

In general,

$$\mathbb{P}(X_1+X_2=k) = \sum_x p_{X_1}(x)\,p_{X_2}(k-x)$$

<u>Example.</u> X_1, X_2 indep't $\quad X_i \sim Poin(\lambda_i) \quad i=1,2$.

$\mathbb{P}(X_1+X_2=k) = 0$, unless $k=0,1,2,\ldots$ in which case,

$$\begin{aligned} \mathbb{P}(X_1+X_2=k) &= \sum_x p_{X_1}(x)\,p_{X_2}(k-x) \\ &= \sum_{x=0}^{\infty} \frac{e^{-\lambda_1}\lambda_1^x}{x!}\,p_{X_2}(k-x) \\ &= \sum_{x=0}^{k} \frac{e^{-\lambda_1}\lambda_1^x}{x!}\,\frac{e^{-\lambda_2}\lambda_2^{k-x}}{(k-x)!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{k!}\sum_{x=0}^{k}\binom{k}{x}\lambda_1^x\lambda_2^{k-x} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{k!}(\lambda_1+\lambda_2)^k \qquad \text{(Binomial Th}^m\text{)} \end{aligned}$$

$\Rightarrow \quad X_1+X_2 \sim Poin(\lambda_1+\lambda_2)$.

$\Rightarrow$ If $X_i \sim Poin(\lambda_i)$ indep't., $\sum_{j=1}^{n} X_j \sim Poin(\sum_1^n \lambda_j)$.

Ex. X_1, X_2 indept, $\quad X_i \sim \mathrm{Bin}(n_i, p) \quad i=1,2 \quad$ (same p)

X_1+X_2 = # of success in n_1+n_2 indept trials $\quad P(S)=p$

$\Rightarrow \quad X_1+X_2 \sim \mathrm{Bin}(n_1+n_2, p)$. More generally, $\sum_1^N X_i \sim \mathrm{Bin}\left(\sum_1^N n_i, p\right)$.

Another calculation: $\quad \mathbb{P}(X_1+X_2=k) = 0$, unless $k=0,\dots,n_1+n_2$, in which case,

$$\mathbb{P}(X_1+X_2=k) = \sum_x \mathbb{P}(X_1=x)\,\mathbb{P}(X_2=k-x)$$

$$= \sum_{x=0}^{k} \binom{n_1}{x} p^x (1-p)^{n_1-x} \binom{n_2}{k-x} p^{k-x} (1-p)^{n_2-(k-x)}$$

$$= p^k (1-p)^{(n_1+n_2)-k} \underbrace{\sum_{x=0}^{k} \binom{n_1}{x}\binom{n_2}{k-x}}$$

# of ways to pick from n_1+n_2 things k things unordered $= \binom{n_1+n_2}{k}$

$$\mathbb{P}(X_1+X_2=k) = \begin{cases} \binom{n_1+n_2}{k} p^k (1-p)^{n_1+n_2-k} & , k=0,\dots,n_1+n_2 \\ 0 & , \text{o/w} \end{cases}$$