6 Sums of Independent Random Variables

Suppose that X and Y are independent and continuous with $f_x \& f_y$.
Distribution of X + Y ?

$F_{X+Y}(\mathrm{a}) = \mathrm{P}\{\mathrm{X} + \mathrm{Y} \le \mathrm{a}\}$ called the 'Convolution' of the distributions of $F_x \& F_y$

$$= \int\int_{x+y\le a} f_X(x) f_Y(y) dx dy$$

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

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

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

Differentiate

$$f_{X+Y}(a) = \frac{d}{da}\int_{-\infty}^{\infty} F_X(a-y) f_Y(y) dy$$

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

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

Example .

X, Y: Independent and Uniform on (0, 1)

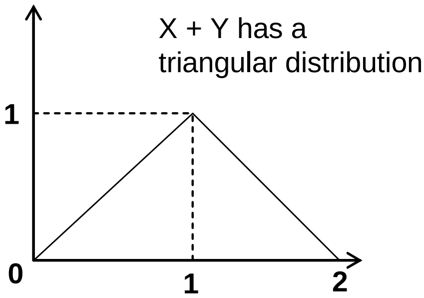
Solution.

$$f_X(a) = f_Y(a) = \begin{cases} 1 & 0 < a < 1 \\ 0 & \text{else} \end{cases}$$

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

$$0 \leq a \leq 1 \quad \Rightarrow \quad f_{X+Y}(a) = \int_0^a \text{dy} = a$$

$$0 < a < 2 \quad \Rightarrow \quad f_{X+Y}(a) = \int_{a-1}^{1} \text{dy} = 2 \text{ - } a$$



Recall Gamma Distribution

$$f(y) = \frac{\lambda e^{-\lambda y}(\lambda y)^{t-1}}{\Gamma(t)} \qquad 0 < y < \infty$$

Proposition

If X and Y are independent gamma random variables with (s, λ) and (t, λ), then X + Y is a gamma random variable with $(s+t, \lambda)$.

Proof.

$$f_{X+Y}(a) = \frac{1}{\Gamma(s)\Gamma(t)} \int_0^a \lambda e^{-\lambda(a-y)}(\lambda(a-y))^{s-1}\lambda e^{-\lambda y}(\lambda y)^{t-1}dy$$

$$= Ke^{-\lambda a}\int_0^a (a-y)^{s-1}\, y^{t-1}dy$$

as, $x = y/a$

$$= Ke^{-\lambda a}a^{s+t-1}\int_0^1 (1-x)^{s-1}\, x^{t-1}dy$$

$$= \text{C } e^{-\lambda a}\, a^{s+t-1}$$

$$\int_{-\infty}^{\infty} f_{X+Y}(a)da = 0 \qquad \Rightarrow \qquad f_{X+Y}(a) = \frac{\lambda e^{-\lambda a}(\lambda a)^{s+t-1}}{\Gamma(s+t)}$$

$\therefore$ For a fixed value of λ, gamma distribution is closed under convolution.

• If $Z_1, Z_2, ..., Z_n$ are independent standard normal random variables, then

$$\text{Y} \equiv \sum_{i=1}^{n} Z_i^2$$

is called the χ^2 (chi-squared) distribution with n degree of freedom.

• When n=1, Y = Z_i^2,

$$\begin{aligned} f_{z^2}(y) &= \frac{1}{2\sqrt{y}} \left[f_z(\sqrt{y}) + f_z(-\sqrt{y}) \right] \\ &= \frac{1}{2\sqrt{y}} \frac{2}{2\sqrt{2\pi}} e^{-\frac{y}{2}} \\ &= \frac{1}{\sqrt{\pi}} \cdot \frac{1}{2} e^{-\frac{y}{2}} \left(\frac{y}{2}\right)^{\frac{1}{2}-1} \end{aligned}$$

Gamma distribution with $(\frac{1}{2}, \frac{1}{2})$

• The χ^2 distribution with n degree of freedom is just the gamma distribution with $(\frac{n}{2}, \frac{1}{2})$.

$$\begin{aligned} \therefore\ f_{\chi^2}(y) &= \frac{\frac{1}{2} e^{-\frac{y}{2}} \left(\frac{y}{2}\right)^{\frac{1}{2}-1}}{\Gamma(\frac{n}{2})} \qquad \text{y} > 0 \\ &= \frac{e^{-\frac{y}{2}} y^{\frac{1}{2}-1}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \qquad \text{y} > 0 \end{aligned}$$

• $\Gamma(\frac{n}{2}) = \begin{cases} \left(\frac{n}{2}-1\right)! & \text{n even} \\ \left(\frac{n}{2}-1\right)\Gamma\left(\frac{n}{2}-1\right) = ... = \left(\frac{n}{2}-1\right) ...... \frac{3}{2}\,\frac{1}{2}\,\Gamma(\frac{n}{2}) & \text{n odd} \end{cases}$

• χ^2 distribution often arises as the distribution of the square of the error involved when attempting to hit a target in n dimensional space, when the coordinate error are taken to be independent unit normal random variables.

Proposition

$X_i,\ i = 1, ..., n$: independent random variables $\sim N(\mu_i, {\sigma_i}^2)$.

$$\text{Then} \qquad \sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i,\ \sum_{i=1}^{n} {\sigma_i}^2\right).$$

Proof.

Let X, Y : independent

$$X \sim N(0, {\sigma_i}^2)$$

$$Y \sim N(0,1)$$

Then,

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

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

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

$$= \frac{1}{\sqrt{2\pi}\ \sigma}\ \exp\left[-\frac{a^2}{2\sigma^2}\right]\ \exp\left[\frac{a}{2\sigma^2(1+\sigma^2)}\right]\cdot\int_{-\infty}^{\infty}\exp\left[-C\left(y-\frac{a}{1+\sigma^2}\right)^2\right]dy$$

$$= \ \mathrm{C}\ \exp\left[-\frac{a^2}{2(1+\sigma^2)}\right]$$

$$\therefore \mathrm{X} + \mathrm{Y}\ \ \sim N(0, 1+\sigma^2).$$

Now suppose X_1, X_2 indept, $X_i \sim N(\mu_i, {\sigma_i}^2)$

Then,

$$X_1 + X_2 = \sigma_2\left(\underbrace{\underbrace{\frac{X_1-\mu_1}{\sigma_2}}_{normal\ with\ mean\ 0,\ var\ \frac{{\sigma_1}^2}{{\sigma_2}^2}} + \underbrace{\frac{X_2-\mu_2}{\sigma_2}}_{normal\ with\ mean\ 0,\ var\ 1}}_{N\left(0,\ 1+\frac{\sigma_1^2}{\sigma_2^2}\right)}\right) + \ \mu_1+\mu_2\ .$$

$$\therefore \qquad X_1 + X_2\ \sim\ N\left(\mu_1+\mu_2,\ \sigma_2^2\left(1+\frac{\sigma_1^2}{\sigma_2^2}\right)\right)$$

By Induction,

$$\sum_{i=1}^{n} X_i \sim N\left(\sum \mu_i, \sum {\sigma_i}^2\right).$$

• Y is said to be a lognormal random variable with parameters μ & σ if log (Y) is a normal random variable with mean μ and var σ^2, i.e. Y is lognormal if Y = e^X, where X is a normal random variable.

Example .

X, Y : indept Poisson random variables with λ_1, λ_1

X + Y ?

Solution.

$$\text{P}\{\text{X + Y = n}\} = \sum_{K=0}^{n} \text{P}\{\text{X = K, Y = n - K}\}$$

$$= \sum \text{P}\{\text{X = K}\}\ \text{P}\{\text{Y = n - K}\}$$

$$= \sum e^{-\lambda_1} \frac{\lambda_1^K}{K!}\ e^{-\lambda_2} \frac{\lambda_2^{n-K}}{(n-K)!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{K=0}^{n} \frac{n!}{K!(n-K)!}\ \lambda_1^K\ \lambda_2^{n-K}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{n!}\ (\lambda_1 +_2)^n$$

$\therefore$ X + Y has a Poisson distribution with parameter $\lambda_1 + \lambda_2$.

Example .

X, Y : indept binomial (n, p), (m, p).

Solution.
Even without any computation, we can see that X + Y is binomial with $(n+m, p)$.
why ?
X: # of successes in n indept trials, each with success probability p.
Y: # of successes in m indept trials, each with success probability p.
X + Y: # of successes in n+m indept trials, each with success probability p.

Lets check

$$\text{P}\{\text{X + Y = K}\} = \sum_{i=0}^{n} \text{P}\{\text{X = i, Y = K - i}\}$$

$$= \sum \text{P}\{\text{X = i}\}\ \text{P}\{\text{Y = K - i}\}$$

$$= \sum \left(\begin{array}{c} n \\ i \end{array} \right) p^i q^{n-i} \left(\begin{array}{c} m \\ K-i \end{array} \right) p^{k-i} q^{m-k+i}$$

$$= p^K q^{n+m-k} \underbrace{\sum \left(\begin{array}{c} n \\ i \end{array} \right) \left(\begin{array}{c} m \\ K-i \end{array} \right)}_{\left(\begin{array}{c} n+m \\ K \end{array} \right)}$$