7 Mar 2022

Gaussian distributions

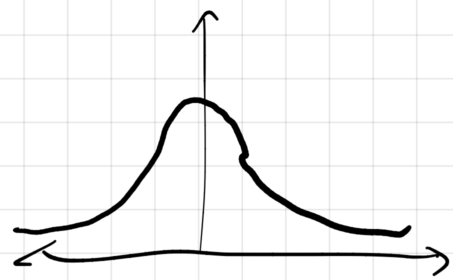
If X is a random variable with continuous, strictly increasing CDF, F, then one way to sample from distrib of X is to draw Y uniformly from $[0,1]$ and put $X = F^{-1}(Y)$.

E.g. if X is exponential with rate r, i.e.

$$\Pr(X > t) = e^{-rt} \quad \forall t$$

then $X = F^{-1}(Y) = \frac{1}{r} \ln\left(\frac{1}{1-Y}\right)$.

The Normal Distribution $\mathcal{N}(0,1)$.



Density is $f(x) = \frac{1}{Z} e^{-\frac{1}{2}x^2}$ where Z is a normalizing constant to make $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

Bad news: the CDF $F(x) = \int_{-\infty}^{x} f(y)\,dy$ has no closed form expression.

If (X,Y) are independent $\mathcal{N}(0,1)$ samples, their density is $f(x,y) = \frac{1}{Z^2} e^{-\frac{1}{2}(x^2+y^2)}$

$$[f(x)\,dt + o(dt)] \cdot [f(y)\,dt + o(dt)]$$
$$= f(x)f(y)\,dt^2 + o(dt^2)$$

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \frac{1}{Z^2}\iint e^{-\frac{1}{2}(x^2+y^2)}\,dx\,dy$$

$$= \frac{1}{Z^2}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}r^2}\, r\,dr\,d\theta$$

$$= \frac{2\pi}{Z^2}\int_0^{\infty} r e^{-\frac{1}{2}r^2}\,dr \qquad \text{Subst. } u = \tfrac{1}{2}r^2,\ du = r\,dr$$

$$= \frac{2\pi}{Z^2} \int_0^\infty e^{-u}\, du$$

$$= \frac{2\pi}{Z^2} \Rightarrow Z = \sqrt{2\pi}$$

$N(0,1)$ has density $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$.

If X, Y are indep. $N(0,1)$ and $(X,Y) = (R,\Theta)_{polar}$

then Θ is uniform on $[0, 2\pi]$

R is distrib on $[0,\infty)$ with density $g(r) = re^{-\frac{1}{2}r^2}$

and indep of Θ.

CDF of R is $G(r) = \int_0^r se^{-\frac{1}{2}s^2} ds$ $\quad u = \frac{1}{2}s^2$

$du = s\, ds$

$$= \int_0^{\frac{1}{2}r^2} e^{-u}\, du$$

$$= 1 - e^{-\frac{1}{2}r^2}$$

Sample R by drawing Y from Unif $[0,1]$.

$$R = \sqrt{2 \ln\left(\frac{1}{1-Y}\right)}.$$

Procedure for sampling $N(0,1)$:

1. Draw $\Theta \sim$ Unif $[0, 2\pi]$
2. Draw $Y \sim$ Unif $[0,1]$
3. $R = \sqrt{2 \ln\left(\frac{1}{1-Y}\right)}$
4. Output $R\cos(\Theta)$.

Observe $\Pr(R < r) = 1 - e^{-\frac{1}{2}r^2}$

$\therefore$ $$\Pr(R^2 < t) = 1 - e^{-\frac{1}{2}t}$$

$\therefore$ R^2 is exponentially distrib with ~~rate~~ $\frac{1}{2}$.

$\therefore$ $$E[R^2] = 2.$$

Remember $R^2 = X^2 + Y^2$ and each of X, Y is $N(\overset{\text{mean}}{0}, \overset{\text{var.}}{1})$.

$\therefore$ $$E[X^2] = 1$$

Since $E[X] = 0$, this means $\mathrm{Var}(X) = 1$.

Why important?

<u>Central Limit Theorem</u>: If $X_1, X_2, X_3, \ldots$ is an infinite seq of identically distributed random variables each with mean μ and variance $\sigma^2 < \infty$, then

$$\frac{\sqrt{n}}{\sigma} \cdot \left(\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right) \xrightarrow{d} N(0,1)$$

If you average n random numbers with mean μ you'll get something close to μ but it'll differ by about $\frac{\sigma}{\sqrt{n}}$.

The Multivariate Normal Distribution $N(0, \mathbb{1})$ is the distrib of d indep. $N(0,1)$ random variables.

$\leftarrow$ d×d identity

density $f(x_1, \dots, x_d) = \left(\frac{1}{\sqrt{2\pi}}\right)^d \cdot e^{-\frac{1}{2}(x_1^2 + x_2^2 + \dots + x_d^2)}$

depends only on $\|(x_1, \dots, x_d)\|_2$.

$\therefore$ rotationally invariant.

If $X = (X_1, \dots, X_d)$ is a sample from $N(0, \mathbb{1})$ and Q is orthogonal matrix, QX is also distributed according to $N(0, \mathbb{1})$.

These 2 properties:

(1) the co-ordinates of X are independent rand vars.

(2) rotation invariant

are important.

<u>ILLUSTRATION</u>: Say X_1, X_2 are indep. $N(0,1)$.

What is the distrib of $X_1 + X_2$?

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$ is an orthogonal matrix

$Q\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ has $N(0, \mathbb{1})$ distribution.

Its first coordinate is $\frac{1}{\sqrt{2}}(X_1 + X_2)$.

$\frac{X_1 + X_2}{\sqrt{2}}$ is $N(0,1)$ distributed $\Rightarrow X_1 + X_2$ is $N(0,2)$.