1

Lecture 13: Normal Distribution

Universality of Uniform

Let F be a continuous, strictly increasing CDF.

Then $X = F^{-1}(U) \sim F$ if $U \sim \text{Unif}(0,1)$.

Started with uniform distribution computed $F^{-1}(u)$ and we claimed that it has CDF F.

Also: (started with X and we don't have a uniform yet) we just have X which has CDF F)

i.e.,

if $X \sim F$, then $F(X) \sim \text{Unif}(0,1)$

perfectly valid random variable

F is just a function

function of a random variable is a random variable

plug X into its own CDF

$F(x) = P(X \leq x)$

$F(X) = P(X \leq X) = 1$ ✗ (wrong)

Ex

$F(x) = 1 - e^{-x}, x > 0$ (Important CDF, exponential distribution).

Interpretation would be

$F(X) = 1 - e^{-X}$

$F(X) \sim \text{Unif}(0,1)$ [Useful in Stats 111, statistical Inference]

→ The interpretation of this is — we first evaluate the function F as a function (something squared or whatever). Write it as a function, then replace x by X.

Example (Useful in simulation)

Let $F(x) = 1 - e^{-x}, x > 0$ ($\text{expo}(1)$), $U \sim \text{Unif}(0,1)$, simulate $X \sim F$.

(By universality theorem) $F^{-1}(u) = -\ln(1-u) \Rightarrow F^{-1}(U) = -\ln(1-U) \sim F$

If we are doing it on a computer and we want 10 random draws from this distribution, we just generate 10 i.i.d. uniforms ~~and~~ (U) and just compute as above 10 times and then we will have 10 i.i.d. random draws from the distribution.

$(1-U) \sim \text{Unif}(0,1)$ (Symmetry of Unif)

$a + bU$ is Unif on some interval.

(0,1)

→ linear transformation

Note: Non-linear usually leads to non Unif.

Independence of random variables $X_1, \ldots, X_n$

$X_1, X_2, \ldots, X_n$ are independent if

$$\underbrace{P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n)}_{\text{Joint CDF}}$$
$$= P(X_1 \le x_1) \cdot P(X_2 \le x_2) \cdots P(X_n \le x_n)$$

individual CDF's

for all $x_1, x_2, \ldots, x_n$

Joint CDF → CDF of all random variables considered jointly as one probability think

Discrete Case

$$\underbrace{P(X_1 = x_1, \ldots, X_n = x_n)}_{\text{Joint PMF}} = P(X_1 = x_1) \cdots P(X_n = x_n)$$

Individual PMF's

Knowing any collections, any subcollection of these X's, Knowing their values tells us nothing about the other ones.

<u>Note</u> :- This is stronger than just pairwise independence. Pairwise independence says if we know one random variable, it doesnot tell us ~~anything~~ information about any other one random variable.
Full independence means knowing any of them, any collection of them, tells us nothing about ones that we don't know, no information whatsoever.

<u>Example</u> pairwise independence doesnot imply independence.

$$X_1, X_2 \sim \text{Bern}(1/2) \text{ i.i.d.}, \quad X_3 = \begin{cases} 1, & \text{if } X_1 = X_2 \\ 0, & \text{otherwise} \end{cases}$$

i.i.d. coin flips

These are pairwise independent, not independent.

indicator r.v.

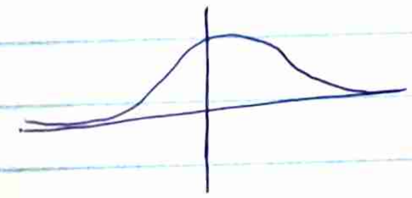
i.e., if two pennies match

- They are not independence because if we know X_1 and X_2, then we immediately know X_3 as X_3 is a function of X_1 and X_2. So, not only does it gives us information, it also gives us total information.
- However, just knowing X_1 tells us nothing about X_2 because those are, we assumed, ~~are~~ independent. Just knowing X_1 tells us nothing about X_3 because it's still 50-50. Similarly, X_2 is independent of X_3. ~~SOT~~
So, they are pairwise independent, but they are not independent.

Pairwise independence isn't enough, in general, just to have independence.

<u>Normal distribution / Gaussian distribution</u>

<u>Central Limit theorem</u> — If we add up a bunch of i.i.d. random variables, the distribution is going to look like a normal distribution.
→ could be either discrete or continuous



$N(0,1)$ — standard Normal dist. notation (mean 0, variance 1)

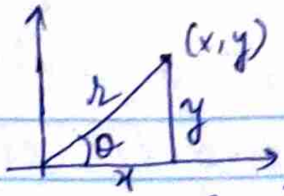
PDF: $f(z) = c e^{-z^2/2}$, c is normalizing constant

<u>Value of c</u>

$$\int_{-\infty}^{\infty} e^{-z^2/2}\,dz \Rightarrow \int_{-\infty}^{\infty} e^{-z^2/2}\,dz \int_{-\infty}^{\infty} e^{-z^2/2}\,dz$$

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx \int_{-\infty}^{\infty} e^{-y^2/2}\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\,dx\,dy$$



$r^2 = x^2+y^2$

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

r → Jacobian

$u = r^2/2$
$du = r\,dr$

$$= \int_0^{2\pi}\Bigg(\underbrace{\int_0^{\infty} e^{-u}\,du}_{1}\Bigg)d\theta$$

$$= 2\pi$$

$$\int_{-\infty}^{\infty} e^{-z^2/2}\,dz = \sqrt{2\pi}$$

$\therefore\ c = 1/\sqrt{2\pi}$

or, $f(z) = \dfrac{1}{\sqrt{2\pi}} e^{-z^2/2}$

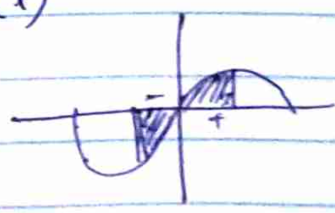
Mean of Standard Normal

$Z \sim N(0,1)$

$EZ = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\infty} z e^{-z^2/2}\,dz = 0$ (by symmetry)

if $g(x)$ is an odd function,
i.e., $g(-x) = -g(x)$

then $\displaystyle\int_{-a}^{a} g(x)\,dx = 0$



Variance

$$Var(Z) = E(Z^2) - (\cancel{EZ})^2$$

$$= E(Z^2)$$

$$= \frac{1}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz}_{\text{evenfunction}} \quad \text{(by LOTUS)}$$

Replace z by z^2

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z^2 e^{-z^2/2} dz$$

$u = z \Rightarrow du = dz$

$dv = ze^{-z^2/2}$

$v = -e^{-z^2/2}$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \underset{u}{z} \cdot \underset{dv}{\underbrace{ze^{-z^2/2}}} dz$$

$$= \frac{2}{\sqrt{2\pi}} \left[\cancel{(uv)}\Big|_0^{\infty} + \int_0^{\infty} e^{-z^2/2} dz \right]$$

$$= \frac{\cancel{2}}{\cancel{\sqrt{2\pi}}} \cdot \frac{1}{\cancel{2}} \cancel{\sqrt{2\pi}}$$

$$= 1$$

Notation

Φ is the standard Normal CDF

phi

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$$

$\Phi(-z) = 1 - \Phi(z)$ (by symmetry)