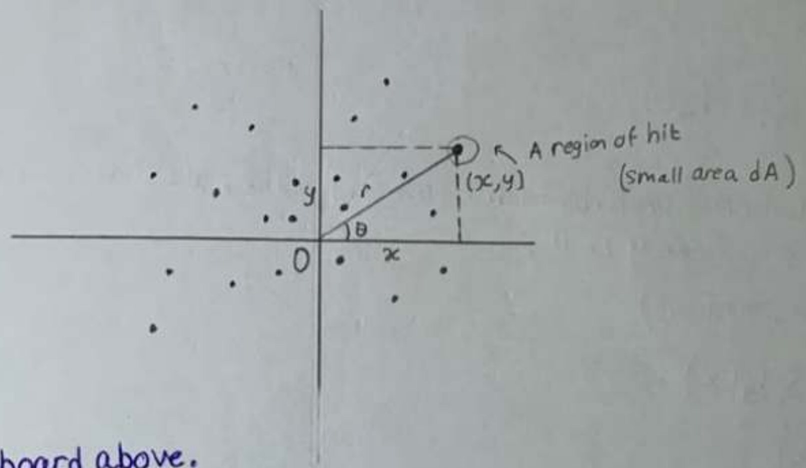
The Normal Distribution



Consider the dartboard above.

- We notice that at a certain reference point O, there are more hits per region and that that value decreases as we move further away from O.

Let us define a probability density function for which when given an input outputs the probability of getting a hit in that region.

The inputs can be defined as follows:

1) Polar Coordinates

Let ρ be a p.d.f

$\Rightarrow \rho(r,\theta)$ where ρ is the distance of the region from O and θ is the angle the region makes with the x-axis.

Note: We assume ρ is rotationally invariant ie. the value of θ does not affect the probability given by ρ

$\Rightarrow \rho(r)$

2) Cartesian Coordinates

Note: we assume that the throws are random.

The random variables X generate values of x for horizontal hit
Y generate values of y for vertical hit

$\Rightarrow$ Let f be a p.d.f

$f_{XY}(x,y) \longrightarrow$ Joint probability distribution

- We also assume that horizontal and vertical hits are statistically independent ie. "How much I miss horizontally does not affect how much I miss vertically".

$\Rightarrow f_{XY}(x,y) = f_X(x)\, f_Y(y) \rightarrow$ Defn of statistical independence

1) The hits are radially symmetric ie. the distributi[on] of hits along any radius (from O) is the same
$\Rightarrow$ Same p.d.f acts on both horizontal & vertical.

$\Rightarrow f_X(x) f_Y(y) = f(x) f(y)$

The probability of getting a hit at a certain region (x,y) is given by the probaility that it hits the x coordinate of the region AND the y-coordinate of the region.

$$\Rightarrow \rho(r) = f_x(x)\, f_y(y)$$

Now, $r = \sqrt{x^2 + y^2}$

$$\Rightarrow \rho\left(\sqrt{x^2+y^2}\right) = f_x(x)\, f_y(y)$$

- To better understand the probability distribution functions we shall work with only one variable. We therefore set y to 0

$$f_y(\overset{y}{0}) = \lambda \text{ (some constant)}$$

$$\Rightarrow \rho\left(\sqrt{x^2+0^2}\right) = \lambda f_x(x)$$

$$\Rightarrow \rho(x) = \lambda f_x(x)$$

We see that ρ is just some scaled version of $f_x(x)$ for all inputs x

ie. for input $\sqrt{x^2+y^2}$

$$\rho\left(\sqrt{x^2+y^2}\right) = \lambda f\left(\sqrt{x^2+y^2}\right)$$

Remember $\rho\left(\sqrt{x^2+y^2}\right) = f(x) f(y)$

$$\Rightarrow f(x) f(y) = \lambda f\left(\sqrt{x^2+y^2}\right)$$

We have eliminated the function with polar arguments, ρ from our equation.

divide by λ^2 *

$$\frac{f(x)}{\lambda} \cdot \frac{f(y)}{\lambda} = \frac{\cancel{\lambda} f\left(\sqrt{x^2+y^2}\right)}{\lambda^{\cancel{2}}}$$

* This was done to better study the nature of f by temporarily getting rid of λ

Let $g(x) = \dfrac{f(x)}{\lambda}$ *

$$\Rightarrow \boxed{g(x)\, g(y) = g\left(\sqrt{x^2+y^2}\right)}$$

Now we need a function g which satisfies the following properties :

g such that : 1) For any given input, the value is first squared
ie. $g_1 : x \mapsto x^2$

2) After squaring the inputs the function must satisfy the following property: $g_2(a)\, g_2(b) = g_2(a+b)$

↳ This is satisfied by <u>exponentiation</u>.

ie. $g_2 : x \mapsto A^x$

$\therefore g = g_2 \circ g_1$

$\Rightarrow g : x \mapsto A^{x^2}$ ⟶ for $g(x)\, g(y) = g(\sqrt{x^2+y^2})$

$A^{x^2} \times A^{y^2} = A^{x^2+y^2}$ ⟶ Validation from law of exponents.

By Convention, let us work with base e

- Via change of base

$$g(x) = e^{\ln(A) x^2}$$

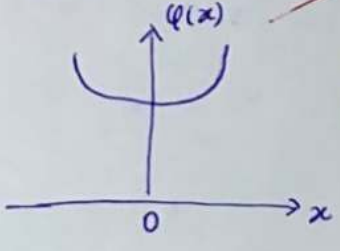
$$\ln(A) = K$$

$$\Rightarrow g(x) = e^{Kx^2}$$

$$\Rightarrow f(x) = \lambda e^{Kx^2}$$

Note : From now on $f(x)$ will be <u>renamed</u> to $\varphi(x)$.

$$\varphi(x) = \lambda e^{kx^2}$$



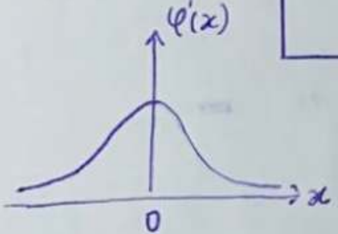
(and centered)
We notice that $\varphi(x)$ is symmetric about the origin 0.
ie. $\varphi(x) = \varphi(-x)$

As we can see, as we integrate $\varphi(x)$ over $(-\infty, \infty)$, the integral blows up. Also, this is contrary to our assumption that the probability of hit decreases as we move away from centre 0.
To remedy this we define $\varphi'(x)$ as follows:

$$\varphi'(x) = \lambda e^{-k^2x^2}$$

Such that the operation of φ on x is still the same (ie. first squares the input and then exponentiate it.) Indeed in that case, Probability decreases as we move away from centre 0.

we introduce a squared value to ensure that the terms in the exponent always remain negative.



- Now, since we are dealing with probabilities we want $P(\text{total}) = 1$

ie. $\int_{-\infty}^{\infty} \varphi'(x)\, dx \stackrel{!}{=} 1$

To do so, we introduce a normalising ~~factor~~ Condition as follows:

$$\int_{-\infty}^{\infty} \lambda e^{-k^2x^2} dx \stackrel{!}{=} 1$$

U-subs

Let $u = kx$

$dx = \frac{1}{k} du$

$\Rightarrow \frac{\lambda}{k} \int_{-\infty}^{\infty} e^{-u^2} du \stackrel{!}{=} 1$

$\sqrt{\pi}$ (Gaussian Integral) *

$\Rightarrow \frac{\lambda\sqrt{\pi}}{k} \stackrel{!}{=} 1$

$\Rightarrow k \stackrel{!}{=} \lambda\sqrt{\pi}$ for $P(\text{total}) = 1$

~~Gaussian Integral~~ *

I/ Gaussian Integral

$I = \int_{-\infty}^{\infty} e^{-x^2} dx \leftarrow$ Even function *

$= 2\int_{0}^{\infty} e^{-x^2} dx$

- We also define I in terms of y

ie. $I = 2\int_0^{\infty} e^{-y^2} dy$

$\Rightarrow I^2 = 4\int_0^{\infty}\int_0^{\infty} e^{-(x^2+y^2)} dy\, dx$

Now let $y = xs \Rightarrow dy = x\, ds$

Changing Bounds

- when $y = 0$

$\Rightarrow 0 = xs \leftarrow$ $s = 0$ makes this valid.

$S = 0$ since x can take any value from 0 to ∞.

- when $y \to \infty$

$\lim_{y\to\infty} (y) = \infty$

$\lim_{s\to ?} (sx) \stackrel{!}{=} \infty$

S must also tend to ∞ as x can take any value $(0, \infty)$, ie $s \to \infty$ makes this valid.

$I^2 = 4\int_0^{\infty}\int_0^{\infty} e^{-(x^2+x^2s^2)} x\, ds\, dx$

$= 4\int_0^{\infty}\int_0^{\infty} x e^{-x^2(1+s^2)} ds\, dx$

From the graph of $f(x,s) = e^{-x^2(1+s^2)}$, it can be seen that $f(x,s)$ is continuous over $\mathbb{R}^2 \Rightarrow$ By Fubini's theorem, order of integration can be exchanged.

$I^2 = 4\int_0^{\infty}\left(\int_0^{\infty} x e^{-x^2(1+s^2)} dx\right) ds$

$= 4\int_0^{\infty}\left[-\frac{1}{2(1+s^2)} \cdot e^{-x^2(1+s^2)}\right|_0^{\infty} ds$

$= 2\int_0^{\infty} \frac{1}{1+s^2} ds$

$= 2\left[\tan^{-1} s\right|_0^{\infty}$

$= 2 \times \frac{\pi}{2} = I^2$

$\Rightarrow \boxed{I = \sqrt{\pi}}$

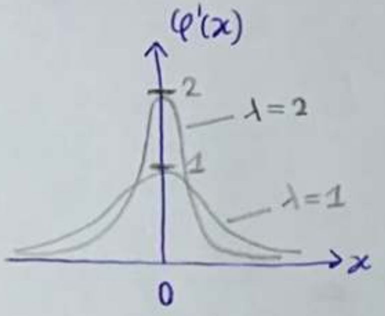
* Note: we made lower bound 0 (NOT $-\infty$) else the limits of s would depend on sign of x.
ie for $(-\infty, \infty)$

when x is -ve
$\lim_{s\to+\infty}(-xs) = -\infty$

when x is +ve
$\lim_{s\to-\infty}(+xs) = -\infty$

$$\Rightarrow \varphi'(x) = \lambda e^{-\lambda^2 \pi x^2}$$

- Now let us look how λ affects $\varphi'(x)$



- We notice as we increase λ
  1) The peak of the curve gets higher
  2) The curve gets "thinner" (ie. more clustered about the mean.)

$\rightarrow$ In Statistics, this can be interpreted as a decrease in Standard deviation (σ).

$$\Rightarrow \boxed{\lambda \propto \frac{1}{\sigma}}$$

Def$\underline{\underline{n}}$ of Variance (σ^2)

For discrete random variables, variance is defined as follows:

$$Var(X) = \sigma^2 = \sum_{i=1}^{n} P_i (x_i - \mu)^2 = \left(\sum_{i=1}^{n} P_i x_i^2\right) - \mu^2$$

$\uparrow$ Gives the probability of each x_i

For our p.d.f we have: i) μ (mean) at 0
ii) $n \rightarrow \infty$
iii) x takes continuous values on $(-\infty, \infty)$

The equation for variance becomes:

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 \varphi'(x)\, dx - \cancel{\mu^2}_{0}$$

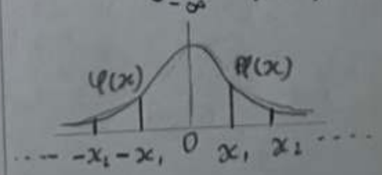
$$\sigma^2 = \int_{-\infty}^{\infty} x^2 \lambda e^{-\lambda^2 \pi x^2}\, dx$$

$$\sigma^2 = \lambda \left[\int_{-\infty}^{\infty} x^2 e^{-\lambda^2 \pi x^2}\, dx\right]$$

$\searrow$ IBP

Note: For D.r.v $\mu = \sum_{i=1}^{n} x_i P_i$

$\rightarrow$ For the above conditions
$\mu = \int_{-\infty}^{\infty} x \varphi'(x)\, dx \longrightarrow$ Also $x e^{-x^2}$ is an odd function and $\int_{-c}^{c}$ (odd function) d... $= 0$



The net sum of the negative values of x and the positive values yield 0 since $\varphi'(x)$ is symmetric about 0

$\Rightarrow \mu = 0$

$$\sigma^2 = \lambda \int_{-\infty}^{\infty} x^2 e^{-\lambda^2 \pi x^2} dx$$

$$= \lambda \int_{-\infty}^{\infty} x \cdot x e^{-\lambda^2 \pi x^2} dx \longrightarrow$$ We know $x e^{-x^2}$ can be "easily" integrated.

$U = x \qquad V = \frac{1}{-2\lambda^2\pi x} \cdot x e^{-\lambda^2\pi x^2} = \frac{-1}{2\pi\lambda^2} e^{-\lambda^2\pi x^2}$

$U' = 1 \qquad V' = x e^{-\lambda^2\pi x^2}$

$$\Rightarrow \sigma^2 = \lambda \left[\cancel{\frac{-1}{2\pi\lambda^2} x e^{-\lambda^2\pi x^2} \Big|_{-\infty}^{\infty}} + \frac{1}{2\pi\lambda^2} \int_{-\infty}^{\infty} e^{-\lambda^2\pi x^2} dx \right]$$

odd function

0^*

∞ in negative exponent term converges to 0 faster than linear term x can diverge (to $\pm\infty$)

More details:

Also

Case: $\int_{-\infty}^{\infty} x e^{-x^2} dx$

$$= \int_{-\infty}^{0} x e^{-x^2} dx + \int_{0}^{\infty} x e^{-x^2} dx$$

$$= \lim_{t \to -\infty} \int_{t}^{0} x e^{-x^2} dx$$

$$= \lim_{t \to -\infty} \left(-\frac{e^{-x^2}}{2} \right) \Big|_{t}^{0}$$

$$= \lim_{t \to -\infty} \left(-\frac{e^{-t^2}}{2} \right) \Big|_{t}^{0}$$

$$= \lim_{t \to -\infty} \left(-\frac{1}{2} + \frac{1}{2} \underbrace{e^{-t^2}}_{\to 0} \right)$$

$$= -\frac{1}{2}$$

$$= \lim_{t \to \infty} \int_{0}^{t} x e^{-x^2} dx$$

$$= \lim_{t \to \infty} \left(-\frac{e^{-x^2}}{2} \right)_{0}^{t}$$

$$= \lim_{t \to \infty} \left(-\underbrace{\frac{e^{-x^2}}{2}}_{\to 0} + \frac{1}{2} \right)$$

$$= \frac{1}{2}$$

$$\Rightarrow -\frac{1}{2} + \frac{1}{2} = 0^*$$

$$\Rightarrow \sigma^2 = \lambda \left[\frac{1}{2\pi\lambda^2} \int_{-\infty}^{\infty} e^{-\lambda^2\pi x^2} dx \right]$$

$$= \frac{1}{2\pi\lambda^2} \underbrace{\int_{-\infty}^{\infty} \lambda e^{-\lambda^2\pi x^2} dx}_{\int_{-\infty}^{\infty} \varphi'(x)\, dx \overset{!}{=} 1} \to 1$$

$$\Rightarrow \sigma^2 = \frac{1}{2\pi\lambda^2}$$

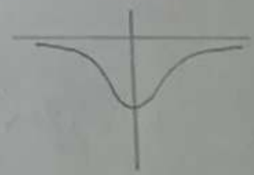
$$\Rightarrow \lambda^2 = \frac{1}{2\pi\sigma^2}$$

$$\varphi'(x) = \lambda e^{-\lambda^2 \pi x^2}$$

$$\lambda = \pm \frac{1}{\sqrt{2\pi}\,\sigma}$$

↳ we do not take the negative value as probability is defined for [0, 1]

Also $-e^{-x^2}$ has shape :

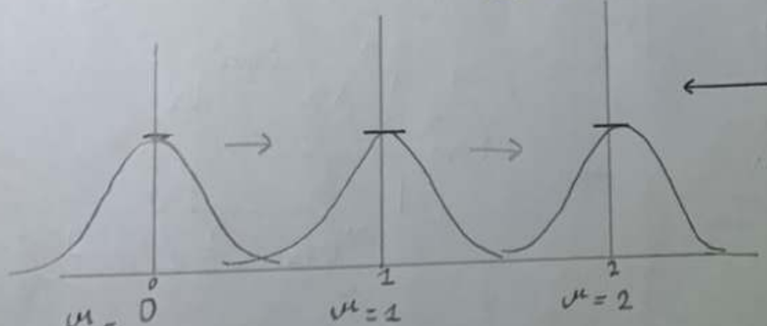


$$\Rightarrow \varphi'(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

Now μ was seen to be 0

To be more general : $\varphi'(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Introducing and varying μ in the exponent shifts the peak (mean) of the curve.



$$\Rightarrow \boxed{\varphi(x \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

Note : $\varphi'(x)$ was renamed to $\varphi(x)$ (without changing any properties of $\varphi'(x)$.)