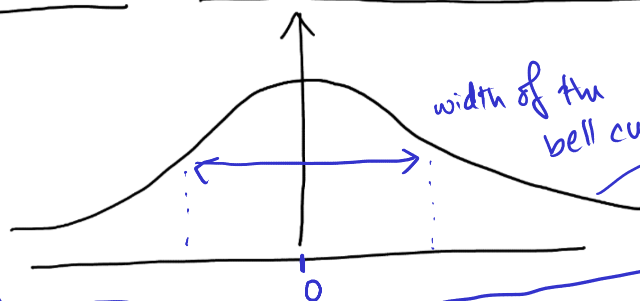
Section 3.5 Gaussian distribution

Normal distribution
"Bell curve."



Normal

variance

$Z \sim N(0,1)$ standard Gaussian if its density

mean

$$\varphi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \qquad -\infty < x < \infty$$

$Z \sim N(\mu, \sigma^2)$

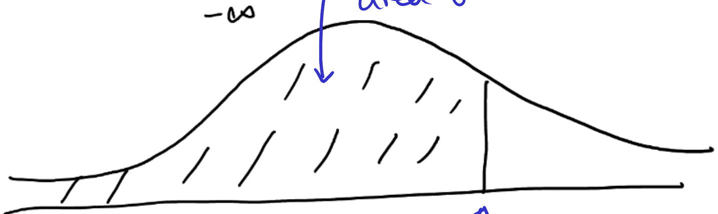
Notation specific to Gaussians

$\varphi \equiv f$ pdf density

$\Phi \equiv F$ cdf

$$\Phi(t) = \mathbb{P}(Z \leq t) = \int_{-\infty}^{t} \varphi(x)\,dx$$

area under the curve



As a "fun" integral lets show $\varphi(x) = \dfrac{e^{-x^2/2}}{\sqrt{2\pi}}$ is a pdf.

1) $\varphi(x) \geq 0$ (exponential fun is +ve)

2) Lets show that $\displaystyle\int_{-\infty}^{\infty} \varphi(x)\,dx = 1 = \Phi(+\infty)$

Equivalent to $I = \displaystyle\int_{-\infty}^{\infty} e^{-x^2/2} = \sqrt{2\pi}$ — ⭐1

"2 replicas"

Here is the trick.

relabeled x to y

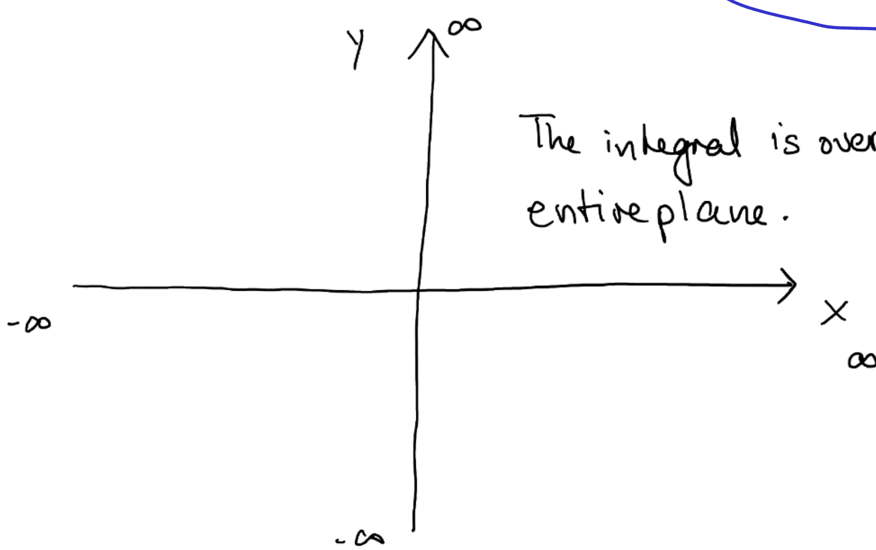
$$I \cdot I = \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \int_{-\infty}^{\infty} e^{-y^2/2}\,dy = \iint \underbrace{e^{-x^2/2}\, e^{-y^2/2}}_{e^{-\left(\frac{x^2}{2}+\frac{y^2}{2}\right)}}\,dx\,dy$$

$\Leftrightarrow$

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

Find Double integral

many, 162



The integral is over the entire plane.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\left(\frac{x^2}{2}+\frac{y^2}{2}\right)\right)dx\,dy = \int_0^{2\pi}\int_0^{\infty} f(r,\theta) \text{ "} r\,dr\,d\theta \text{"}$$

POLL

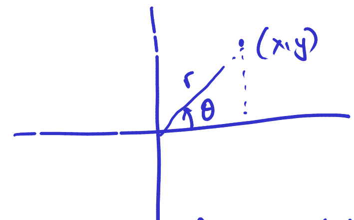
When going from $\iint f(x,y)\,dx\,dy$ to polar coordinates (r,θ). $x = r\cos\theta$ $y = r\sin\theta$ and $dx\,dy$ goes to:

A	B	C
$dr\,d\theta$	$r\,dr\,d\theta$	$\dfrac{dr\,d\theta}{r}$

CHANGING TO POLAR COORDINATES

When you have to change to polar coordinates, $dx\,dy$ becomes $r\,dr\,d\theta$.

Diagram



$$x^2+y^2 = r^2(\cos^2\theta+\sin^2\theta) = r^2$$

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

Matrix

$$\begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\[2ex] \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta - (-r\sin^2\theta) = r(\cos^2\theta+\sin^2\theta) = r$$

what matrix operation is this? Jacobian is a determinant of a matrix

Notice cancellation.

$$dx\,dy = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\[2ex] \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} dr\,d\theta = r\,dr\,d\theta$$

So

$$I^2 = \iint e^{-\frac{(x^2+y^2)}{2}}\,dx\,dy = \int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}}\,r\,dr\,d\theta$$

Compare with ★1 from a few pages ago.

$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}} \, r\,dr\,d\theta = \underbrace{\left(\int_0^{\infty} e^{-r^2/2}\, r\,dr\right)}_{1} \, 2\pi$$

$$\int_0^{2\pi} d\theta = 2\pi$$

$$\int_0^{\infty} e^{-r^2/2}\, r\,dr \qquad \frac{r^2}{2} = u \qquad \frac{2r\,dr}{2} = du$$

$$= \int_0^{\infty} e^{-u}\,du = -e^{-u}\Big|_0^{\infty} = -\left(e^{-\infty} - e^{-0}\right) = (1-0) = 1$$

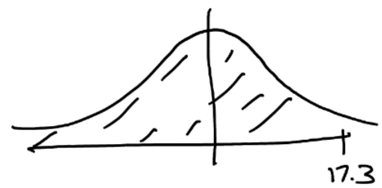
$$I^2 = 2\pi \Rightarrow I = \sqrt{2\pi}$$

$$\Phi(1) = \int_{-\infty}^{1} e^{-x^2/2}\,dx \stackrel{?}{=} \text{I cannot integrate.}$$

In general you cannot compute integrals like this EXPLICITLY.

For example, you cannot

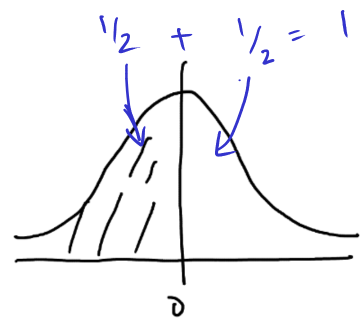
find $\Phi(1)$ $\quad\Phi(3)$ or $\Phi(17.3)$



So you compute things using a ~~calculator~~ and make a table that looks like this:

You can guess this.

	0.00	0.01	0.02	0.03
0.0	.5	0.5040		
0.1				
0.2				
⋮				



$\Phi(0)$ can be found.

$\Phi(0)$ must be $\frac{1}{2}$ by SYMMETRY.

Why do we use tables in this course?

A: It is important to learn how priviledged we are with our fancy computers. Our poor parents used slide rules and abacuses

B: We have to use tables because computers are not allowed on exams

C: There is no good reason. College rules are arbitrary

Notation: $Z \sim \underbrace{N}_{\text{normal / Gaussian}}(0, 1)$

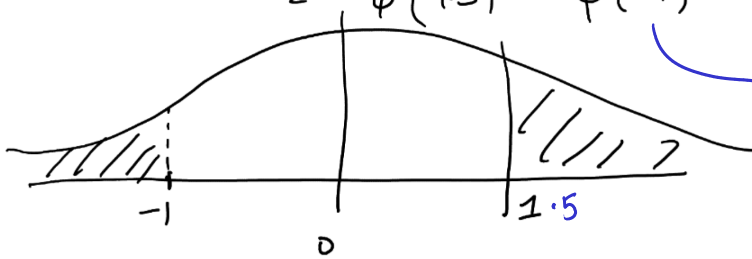
variance = 1

mean 0

Ex 3.42. $Z \sim N(0,1)$ Normal, mean 0 variance 1.

$$\mathbb{P}(-1 < Z \le 1.5) = \mathbb{P}(Z \le 1.5) - \mathbb{P}(Z \le -1)$$

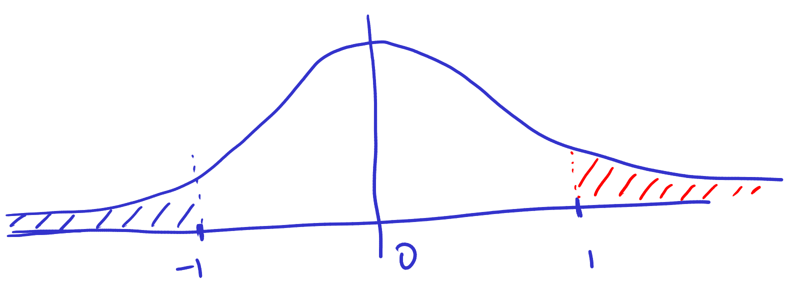
$$= \phi(1.5) - \phi(-1)$$

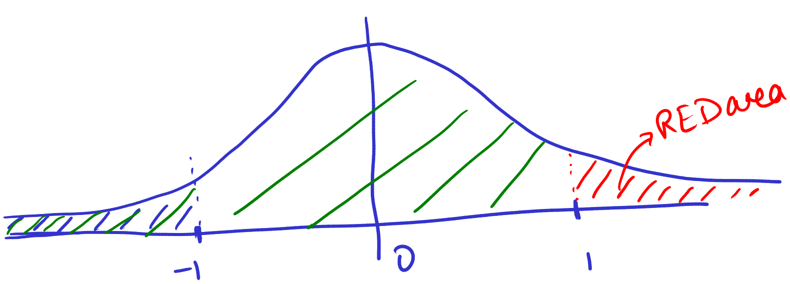


$\Phi(1.5) = 0.9332$

$\Phi(-1) =$

Problem: How do we compute something about $\Phi(-1)$?





$$\text{RED area} = 1 - \underbrace{\text{GREEN}}_{\Phi(1)} = \Phi(-1)$$

$\Rightarrow \quad \Phi(-1) = 1 - \Phi(1) = 1 - 0.8413$

So $\quad \Phi(1.5) - \Phi(-1)$

$= 0.9332 - (1 - 0.8413)$

We claimed that $Z \sim N(0,1)$ has 0 mean and variance 1. Lets prove that.

mean 0

$\sigma^2 = 1$

$$\mathbb{E}[Z] = 0. \quad \int_{-\infty}^{\infty} z \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz = \int_{-\infty}^{0} z \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz + \int_{0}^{\infty} \frac{z e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$$

$$\mathbb{E}[Z] = \int_{-\infty}^{\infty} \frac{z e^{-z^2/2}}{\sqrt{2\pi}} dz = \int_{-\infty}^{\infty} f(z) dz$$

pdf

$$= 0$$

$$f(z) = \frac{z e^{-z^2/2}}{\sqrt{2\pi}}$$

$$f(-z) = \frac{(-z) e^{-\frac{(-z)^2}{2}}}{\sqrt{2\pi}}$$

$f(-z) = -f(z)$ "odd functions"

We can also compute the variance this way:

$$\mathrm{Var}(Z) = E[Z^2] - \cancel{E[Z]}^{\,0\;2} \overset{?}{=} 1$$

2nd "moment"

when in doubt integrate by parts

Enough to compute

$$E[\underbrace{Z^2}_{\text{value}}] = \int_{-\infty}^{\infty} x^2 \underbrace{\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}}_{\text{pdf}} dx$$

$$v = -e^{-x^2/2} \qquad dv = x e^{-x^2/2}$$

$$u = x \qquad du = dx$$

$$= \int_{-\infty}^{\infty} \underbrace{x}_{u} \underbrace{\left(x e^{-x^2/2}\right)}_{dv}$$

$$= uv\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} v\,du = x\left(-\underbrace{e^{-x^2/2}}_{\substack{\text{wins}\\ -\infty \\ -\infty}}\right)\Bigg|_{-\infty}^{\infty} - \underbrace{\int_{-\infty}^{\infty} \left(-e^{-x^2/2}\right) dx}_{\text{we have done this before} = -I}$$

evaluate

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

$$E[Z^2] = \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} = 1 \quad \text{(variance is 1)}$$

NEXT: How do you describe a general normal distribution?

$$\boxed{X \sim N(\mu, \sigma^2)}$$

$\mu = 0,\ \sigma^2 = 1$

$\mu = 7,\ \sigma^2 = 1000$

Mean μ

Variance σ^2.

We claim that general normal distributions can be written as:

$X = \sigma Z + \mu$ when $\boxed{Z \sim N(0,1)}$ Standard normal cdf has values in the appendix

First check mean and variance

$$\mathbb{E}[X] = \sigma \underbrace{\mathbb{E}[Z]}_{0} + \mu = \mu.$$

$$\mathrm{Var}(X) = \sigma^2\, \mathrm{Var}(Z) = \sigma^2$$

$$\mathrm{Var}(\sigma Z + \cancel{\mu}) = \sigma^2\, \mathrm{Var}(Z) = \sigma^2 \cdot 1$$

Ex: $X \sim N(-3, 4)$ $\qquad \mu = -3 \qquad \sigma^2 = 4 \qquad X = \sigma Z + \mu = 2Z - 3$

$$\boxed{\text{Find } P(-1 \le X \le 1.5)} = P(-1 \le 2Z - 3 \le 1.5)$$

$$= P\left(\frac{-1+3}{2} \le \frac{\cancel{2}Z}{\cancel{2}} \le \frac{1.5+3}{2}\right)$$

$$= P\left(\frac{-1+3}{2} \le Z \le \frac{1.5+3}{2}\right)$$

$$P(-1 < X \le 1.5) = P(-1 < X < 1.5) = \Phi\left(\frac{4.5}{2}\right) - \Phi(1)$$

and we can compute this.

$P(X = -1) = P(X = 1.5) = 0$

Here we have used again

$$P(a < X \le b) = F(b) - F(a).$$

worth noting here that

$$P(a \le X \le b) = P(a < X \le b).$$

Since X is continuous. The equality sign in $a < X$ doesn't make a difference

Let $X \sim N(-3, 4)$. Find

1) $P(X > 3.5)$

2) $P(-2.1 < X < -1.9)$

POLL

Q & A style, with voting.

$\mu = -3 \quad \sigma^2 = 4$

$P(X > 3.5)$

$X = \sigma Z + \mu = 2Z - 3$

$\parallel$

$P(2Z - 3 > 3.5)$

$\parallel$

$P(2Z > 6.5) = P(Z > 3.25) = 1 - \Phi(3.25)$

$= 1 - 0.9994$

≈ 0.0006

$P(-2.1 < X < -1.9) = P(-2.1 < 2Z - 3 < -1.9)$

$= P(0.9 < 2Z < 1.1) = P(0.45 < Z < 0.55)$

$= \Phi(0.55) - \Phi(0.45) = 0.7088 - 0.6736$

$=$

$X \sim N(\mu, \sigma)$ $\quad X = \sigma \underset{\uparrow}{Z} + \mu$

Standard normal.

CDF and PDF of general normal

Cdf of X:

$$F_X(t) = P(X \leq t)$$

$$= P(\sigma Z + \mu \leq t) = P(\sigma Z \leq t-\mu)$$

$$= P\left(Z \leq \frac{t-\mu}{\sigma}\right) = F_X(t).$$

$$= \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx = \Phi\left(\frac{t-\mu}{\sigma}\right)$$

Pdf of X:

$$\frac{d F_X(t)}{dt} = f_X(t) = \frac{e^{-\frac{(t-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

Find the cdf of X $\qquad X = \sigma Z + \mu$

$$F_X(t) = P(X \leq t) = P(\sigma Z + \mu \leq t)$$

↑ definition of X

$$= P\left(Z \leq \frac{t-\mu}{\sigma}\right) = \Phi\left(\frac{t-\mu}{\sigma}\right) = \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{e^{-s^2/2}}{\sqrt{2\pi}}\, ds$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{e^{-s^2/2}}{\sqrt{2\pi}}\, ds$$

pdf of X

$$f_X(t) = \frac{d}{dt} F_X(t) = \frac{d}{dt} \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{e^{-s^2/2}}{\sqrt{2\pi}}\, ds$$

$$= \frac{e^{-\left(\frac{t-\mu}{\sigma}\right)^2 \frac{1}{2}}}{\sqrt{2\pi}} \frac{d}{dt}\left(\frac{t-\mu}{\sigma}\right)$$

$$= \frac{e^{-\left(\frac{t-\mu}{\sigma}\right)^2 \frac{1}{2}}}{\sqrt{2\pi}} \frac{1}{\sigma}$$