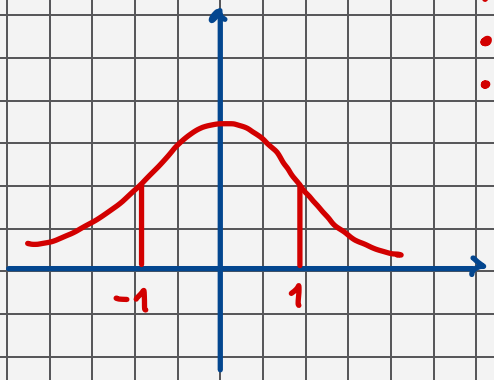
Standard Normal Distribution.

We say that a random variable Z has the

standard normal distribution

If its probability density function (pdf) has the form

$$f_Z(z) = \varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \text{for all } z \in \mathbb{R}$$



- mean / median / mode = 0
- symmetric about the vertical axis, i.e., $f_Z(z) = f_Z(-z)$, $z \in \mathbb{R}$,

i.e., **even**

We write $Z \sim N(0,1)$

The cumulative dist'n f'tion (cdf) of the std normal is

$$N(z) = \Phi(z) = \mathbb{P}[Z \le z]$$

$$= \int_{-\infty}^{z} f_Z(u)\,du = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\,du$$

No Analytic Form!

There are the standard normal tables!

We can use `dnorm` for $\varphi = f_Z$,

and `pnorm` for $\Phi = N$,

and `qnorm` for $\Phi^{-1} = N^{-1}$.

Normal Distributions.

We can completely specify any normal dist'n w/

- its mean μ

and

- its standard deviation σ (or its variance σ^2)

We write $\boxed{X \sim \text{Normal}(\text{mean} = \mu, \text{var} = \sigma^2)}$

If X can be written as a *linear transform* of a *standard normal* Z, i.e.,

$$\boxed{X = \mu + \sigma \cdot Z} \longleftrightarrow \frac{X-\mu}{\sigma} = Z \sim N(0,1),$$

we say that $X \sim \text{Normal}(\text{mean} = \mu, \text{var} = \sigma^2)$

We can check:

- $\mathbb{E}[X] = \mathbb{E}[\mu + \sigma \cdot Z] \overset{\text{Linearity}}{=} \mu + \sigma \underbrace{\mathbb{E}[Z]}_{=0} = \mu$ ✓
- $\text{Var}[X] = \text{Var}[\mu + \sigma \cdot Z]$ (μ: a deterministic shift)
$= \text{Var}[\sigma Z] = \sigma^2 \cdot \underbrace{\text{Var}[Z]}_{=1} = \sigma^2$ ✓

Fact: Consider $X_1 \sim \text{Normal}(\text{mean} = \mu_1, \text{variance} = \sigma_1^2)$

and

$X_2 \sim \text{Normal}(\text{mean} = \mu_2, \text{variance} = \sigma_2^2)$.

Let α_1 and α_2 be real constants.

Assume that X_1 and X_2 are *independent*.

$$\alpha_1 X_1 + \alpha_2 X_2 \sim \text{Normal}\left(\text{mean} = \alpha_1\mu_1 + \alpha_2\mu_2, \text{var} = \alpha_1^2\sigma_1^2 + \alpha_2^2\sigma_2^2\right)$$

The Normal Approximation to the Binomial (de Moivre-Laplace)

Consider a sequence of binomial random variables

$Y_n \sim \text{Binomial}(n = \#\text{ of trials}, p = \text{success probability})$

Then,
- $\mathbb{E}[Y_n] = n \cdot p$
- $\text{Var}[Y_n] = n \cdot p \cdot (1-p) \Rightarrow \text{SD}[Y_n] = \sqrt{n \cdot p \cdot (1-p)}$

$$\frac{Y_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \xrightarrow[n \to \infty]{\mathcal{D}} N(0,1)$$

Usage:
- "Rule of Thumb": n is "large", i.e., $np \geq 10$ and $n(1-p) \geq 10$
- $\mathbb{P}[a \leq Y_n \leq b] =$

$$= \mathbb{P}\left[\frac{a-np}{\sqrt{np(1-p)}} \leq \underbrace{\frac{Y_n - np}{\sqrt{np(1-p)}}}_{\approx N(0,1) \sim Z} \leq \frac{b-np}{\sqrt{np(1-p)}}\right]$$

$$\approx \mathbb{P}\left[\frac{a-np}{\sqrt{np(1-p)}} \leq Z \leq \frac{b-np}{\sqrt{np(1-p)}}\right]$$

$$= N\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - N\left(\frac{a-np}{\sqrt{np(1-p)}}\right)$$

N ... cdf of $N(0,1)$

- In statistics: $Y_n \approx N(\text{mean} = np, \text{var} = np(1-p))$
- In m362k: $\mathbb{P}[Y_n = k] = \mathbb{P}[k \leq Y_n \leq k] \approx 0$



PROBLEM. You toss a fair coin 10,000 times. What is the approximate probability that the number of Heads exceeds the number of Tails by between 200 and 500 (inclusive)?

$\rightarrow$: Y... # of Hs

$n-Y$... # of Ts

$$\mathbb{P}[200 \leq Y-(n-Y) \leq 500] =$$

$$= \mathbb{P}[200 \leq 2Y-10000 \leq 500] =$$

$$= \mathbb{P}[10200 \leq 2Y \leq 10500]$$

$$= \mathbb{P}[5100 \leq Y \leq 5250]$$

$$\mu_Y = n \cdot p = 5000 = n(1-p) \geq 10 \quad \checkmark$$

$$\sigma_Y = \sqrt{5000(1/2)} = \sqrt{2500} = 50$$

$$\mathbb{P}\left[\frac{5100-5000}{50} \leq \boxed{\frac{Y-\mu_Y}{\sigma_Y}} \leq \frac{5250-5000}{50}\right]$$

$\approx N(0,1) \sim Z$

$$\approx \mathbb{P}[2 \leq Z \leq 5] = N(5) - N(2) = 1 - 0.9772 = 0.0228$$

□