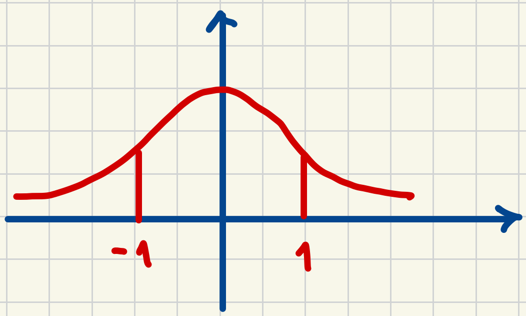
Standard Normal Distribution.

We say that a random variable Z has the **standard normal dist'n** 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.,

$$\boxed{\varphi(z) = \varphi(-z) \text{ for all } z \in \mathbb{R}}$$

i.e., $\boxed{\text{even}}$

The cumulative distribution f'tion (cdf) of the standard normal is

$$N(z) = \Phi(z) = \mathbb{P}[Z \leq z] = \int_{-\infty}^{z} f_Z(u)\,du = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\,du$$

$\boxed{\text{No analytic form!}}$

There are the standard normal tables!

We can use \`dnorm\` for $\varphi = f_Z$ and \`pnorm\` for $\Phi = N$ and \`qnorm\` for $\Phi^{-1} = N^{-1}$ in \`R\`.

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

Normal Distributions.

We can completely specify any normal distribution w/

- its mean μ

and

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

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

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

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

We can check:

- $\mathbb{E}[X] = \mathbb{E}[\mu + \sigma \cdot Z] \overset{\text{linearity of } \mathbb{E}}{=} \mu + \sigma \cdot \underbrace{\mathbb{E}[Z]}_{=0} = \mu$

- $\text{Var}[X] = \text{Var}[\mu + \sigma \cdot Z]$ a deterministic shift

$$= \text{Var}[\sigma \cdot Z] = \sigma^2 \cdot \underbrace{\text{Var}[Z]}_{=1} = \sigma^2$$

Fact. (X_1, X_2) are jointly normal,

then,

$$\alpha_1 X_1 + \alpha_2 X_2 \sim \text{Normal}(\_\_\_\_, \_\_\_\_)$$

In particular,

if X_1 and X_2 are independent,

then,

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

The Normal Approximation to the Binomial

(de Moivre · Laplace)

Consider a sequence of binomial r.v.s

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

Then,

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

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

Usage:

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

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

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

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

$\boxed{N \ldots \text{cdf of } N(0,1)}$

- In statistics: $Y_n \approx N(\text{mean} = np,\ \text{var} = np(1-p))$
- $\boxed{\text{In M362K:}}$ $\mathbb{P}[Y_n = k] = \mathbb{P}[k \leq Y_n \leq k] \approx 0$ ☹

Continuity Correction: