Notation/Terminology:

"Random Sample"

$$X_1, X_2, \ldots, X_n$$

- variables before they are sampled, observed, and "locked in"

- they are assumed to be independent and identically distributed (iid)

random sample = iid

More Notation:

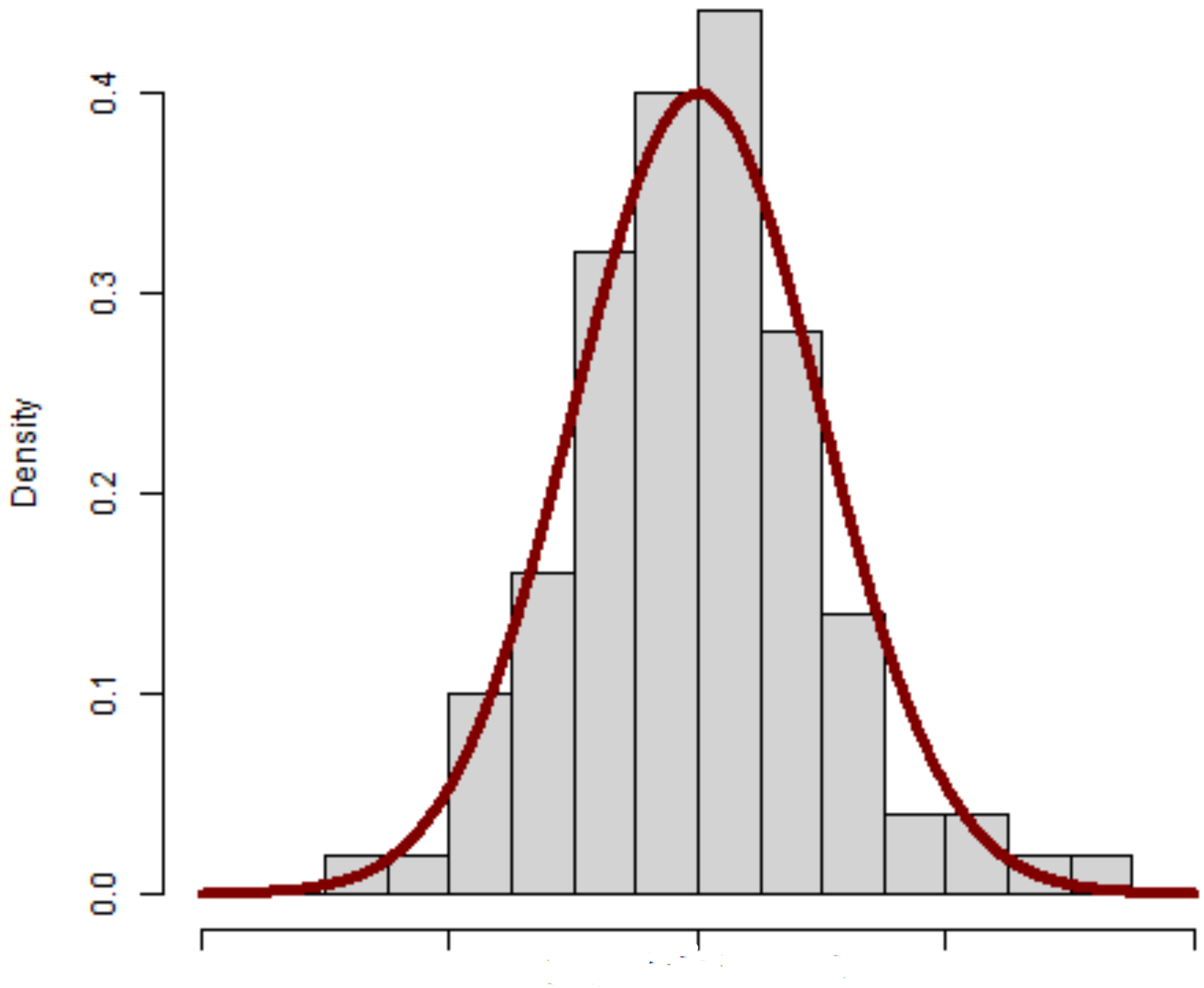
Suppose that $X_1, X_2, \ldots, X_n$ is a random sample from the normal distribution with mean μ and variance σ^2.

We write

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$$

$$E[X_i] = \mu \qquad Var[X_i] = \sigma^2$$

$$X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$$



$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

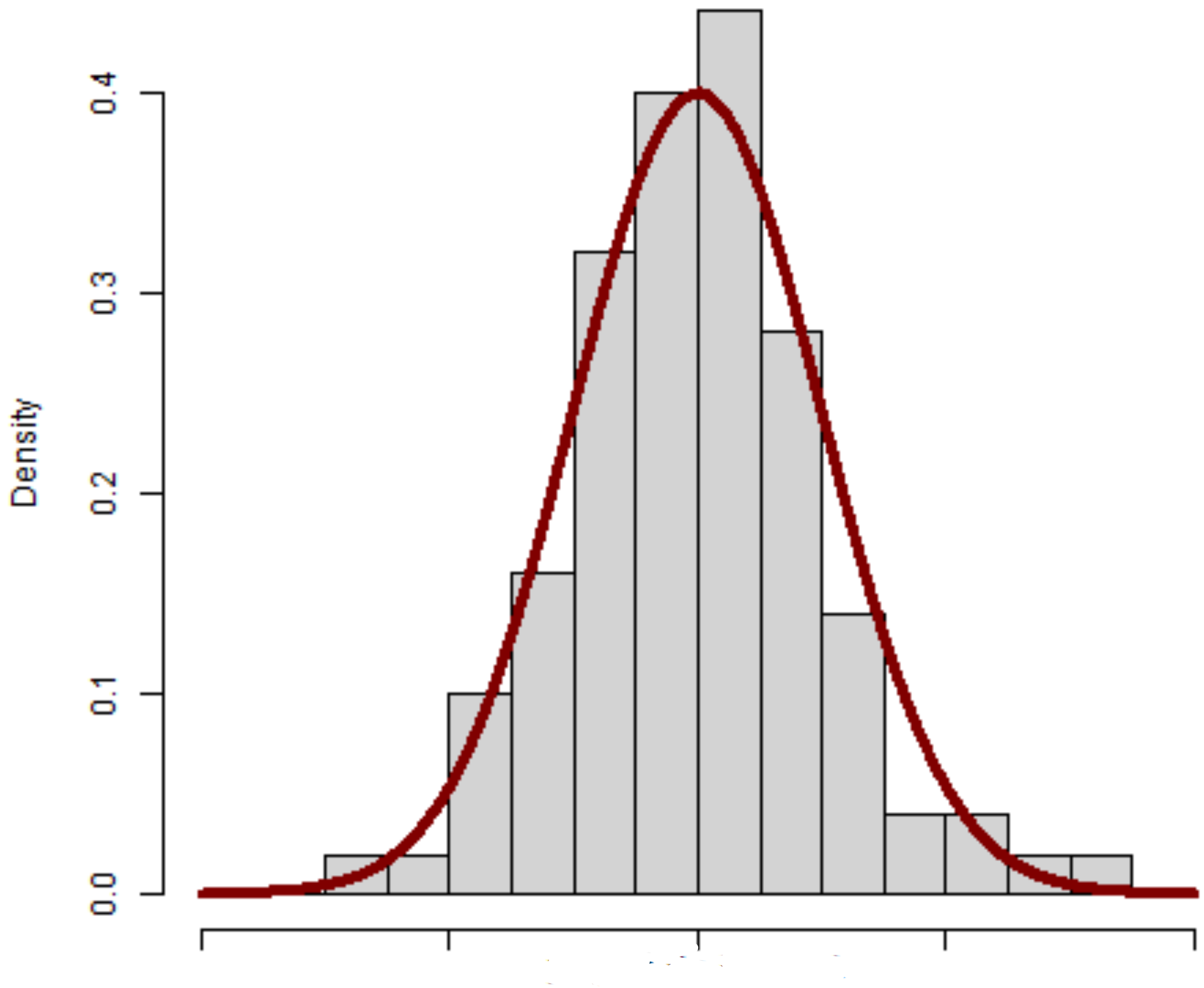
$$\mu = \mathrm{E}[X_i] = \int_{-\infty}^{\infty} x \; f(x)\,dx$$

$$\mathrm{E}[X_i^2] = \int_{-\infty}^{\infty} x^2 \; f(x)\,dx$$

$$\sigma^2 = \mathrm{Var}[X_i] = \mathrm{E}[(X_i - \mu)^2]$$

$$= \mathrm{E}[X_i^2] - (\mathrm{E}[X_i])^2$$

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$$



$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Any linear combination of normal random variables has, again, a normal distribution.

$$a_1 X_1 + a_2 X_2 + \ldots a_n X_n \sim N(?, ?)$$

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i \underbrace{E[X_i]}_{\mu} = \mu \sum_{i=1}^{n} a_i$$

$$\mathrm{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \underbrace{\mathrm{Var}[X_i]}_{\sigma^2} = \sigma^2 \sum_{i=1}^{n} a_i^2$$

Need independence here!

In particular, if

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$$

Then

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Note the smaller variance!

The N(0,1) distribution is known as the **standard normal distribution**.

We typically use the letter Z:

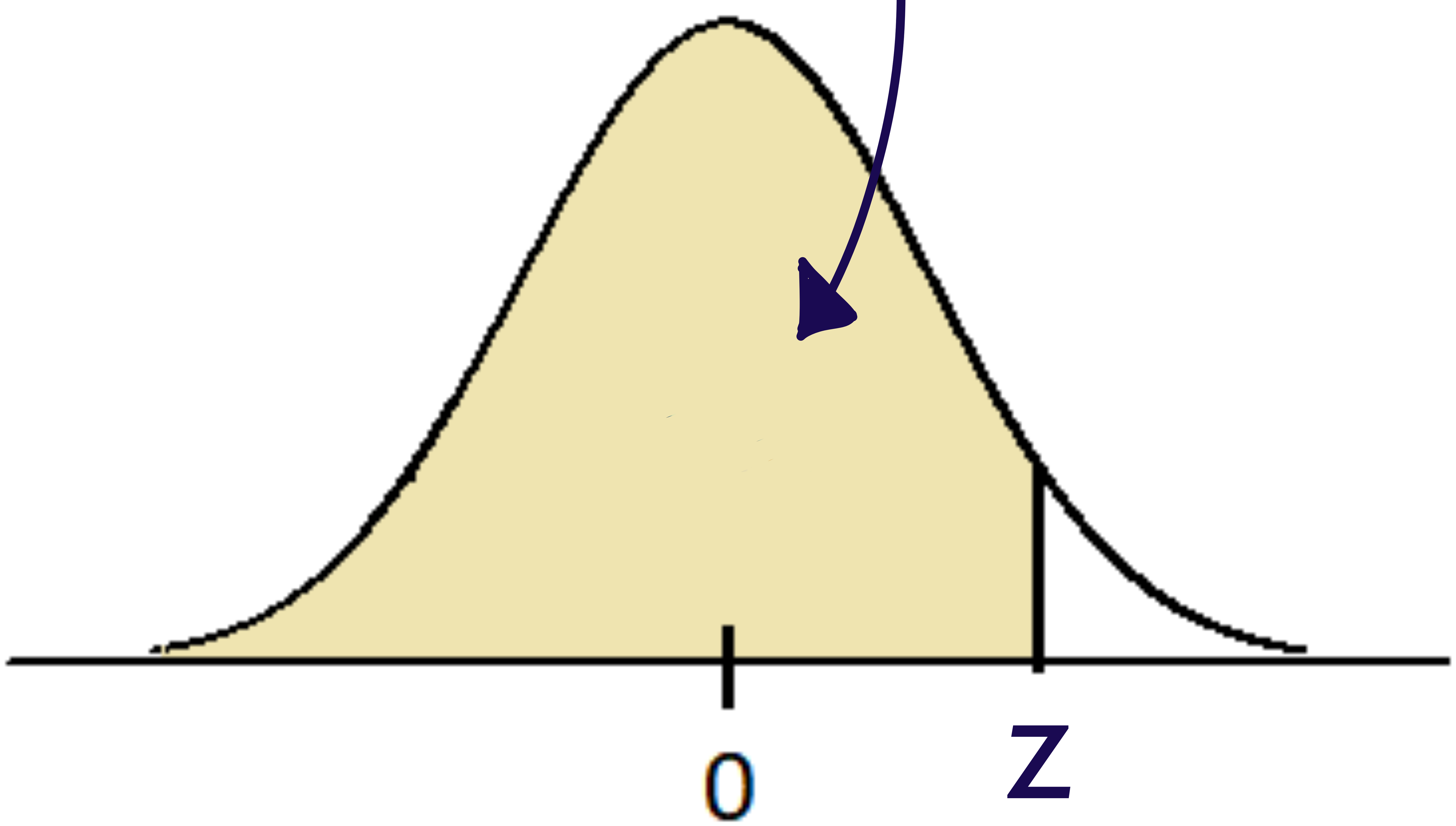
$$Z \sim N(0, 1)$$

The **cumulative distribution function** (cdf)

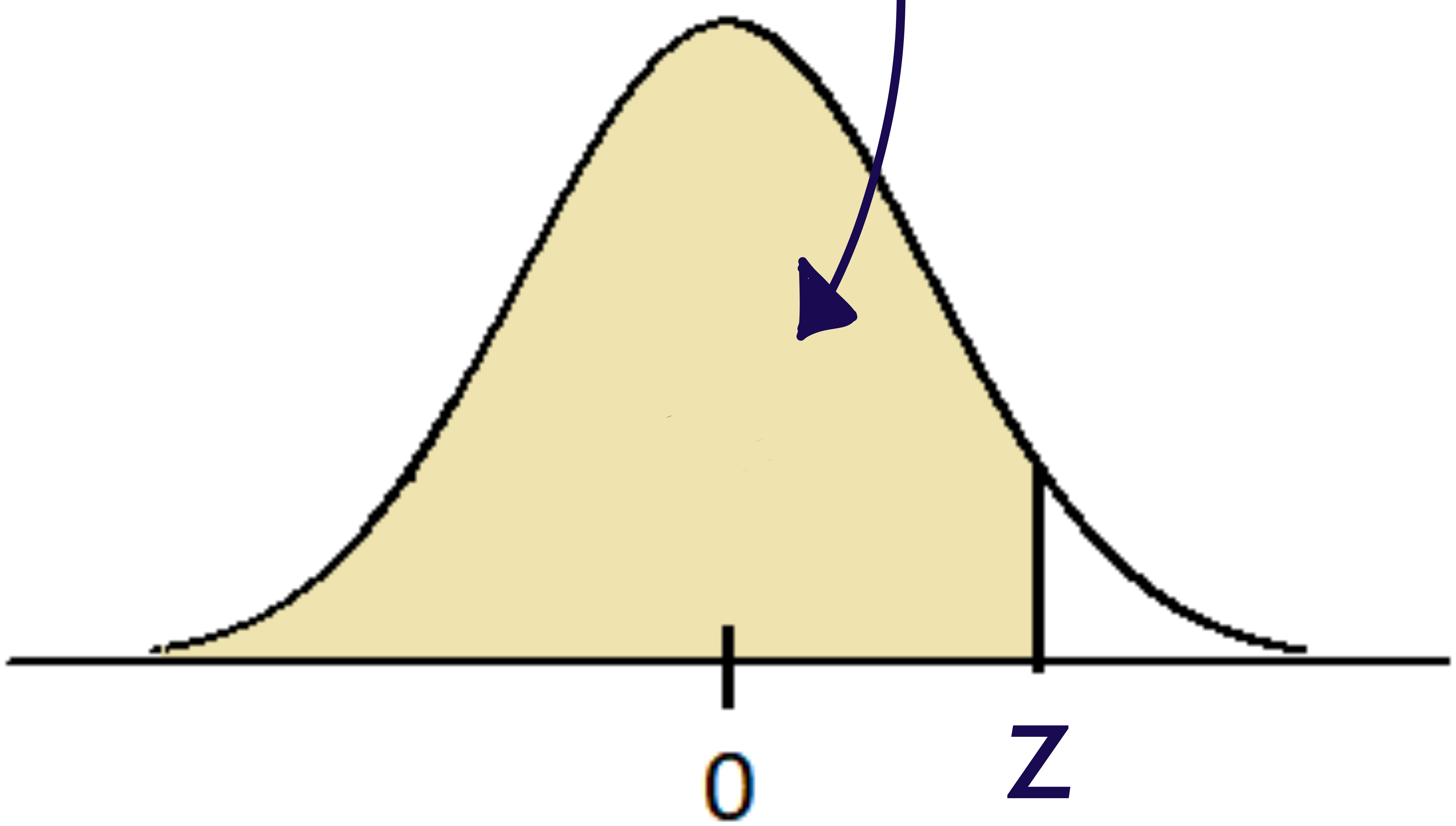
$$\Phi(z) = P(Z \leq z)$$

$$= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

$Z \sim N(0, 1)$ $\Phi(z) = P(Z \leq z)$



$$Z \sim N(0, 1) \qquad \Phi(z) = P(Z \leq z)$$



$$\Phi(2.03) = P(P \leq 2.03) = 0.9788$$

R Code: pnorm(2.03)

Any linear combination of normal random variables has, again, a normal distribution.

$$a_1 X_1 + a_2 X_2 + \dots a_n X_n + b$$

- $X \sim N(\mu, \sigma^2) \;\Rightarrow\; \dfrac{X - \mu}{\sigma} \sim N(0, 1)$
- $Z \sim N(0, 1) \;\Rightarrow\; \sigma Z + \mu \sim N(\mu, \sigma^2)$

Let $X \sim N(2, 3)$.

Then

$$P(X \le 4.1) = P\left(\frac{X-\mu}{\sigma} \le \frac{4.1-2}{\sqrt{3}}\right)$$

$$= P(Z \le 1.21)$$

$$\approx 0.8868$$

R Code: pnorm(1.21)

$$X_1, X_2, \ldots, X_{10} \overset{\text{iid}}{\sim} N(2, 3)$$

$$\overline{X} \sim N(\mu, \sigma^2/n) = N(2, 3/10)$$

$$P(\overline{X} \leq 2.3) = P\left(\frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} \leq \frac{2.3 - 2}{\sqrt{3/10}}\right)$$

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

$$= P(Z \leq 0.5477)$$

$$\approx 0.7081$$

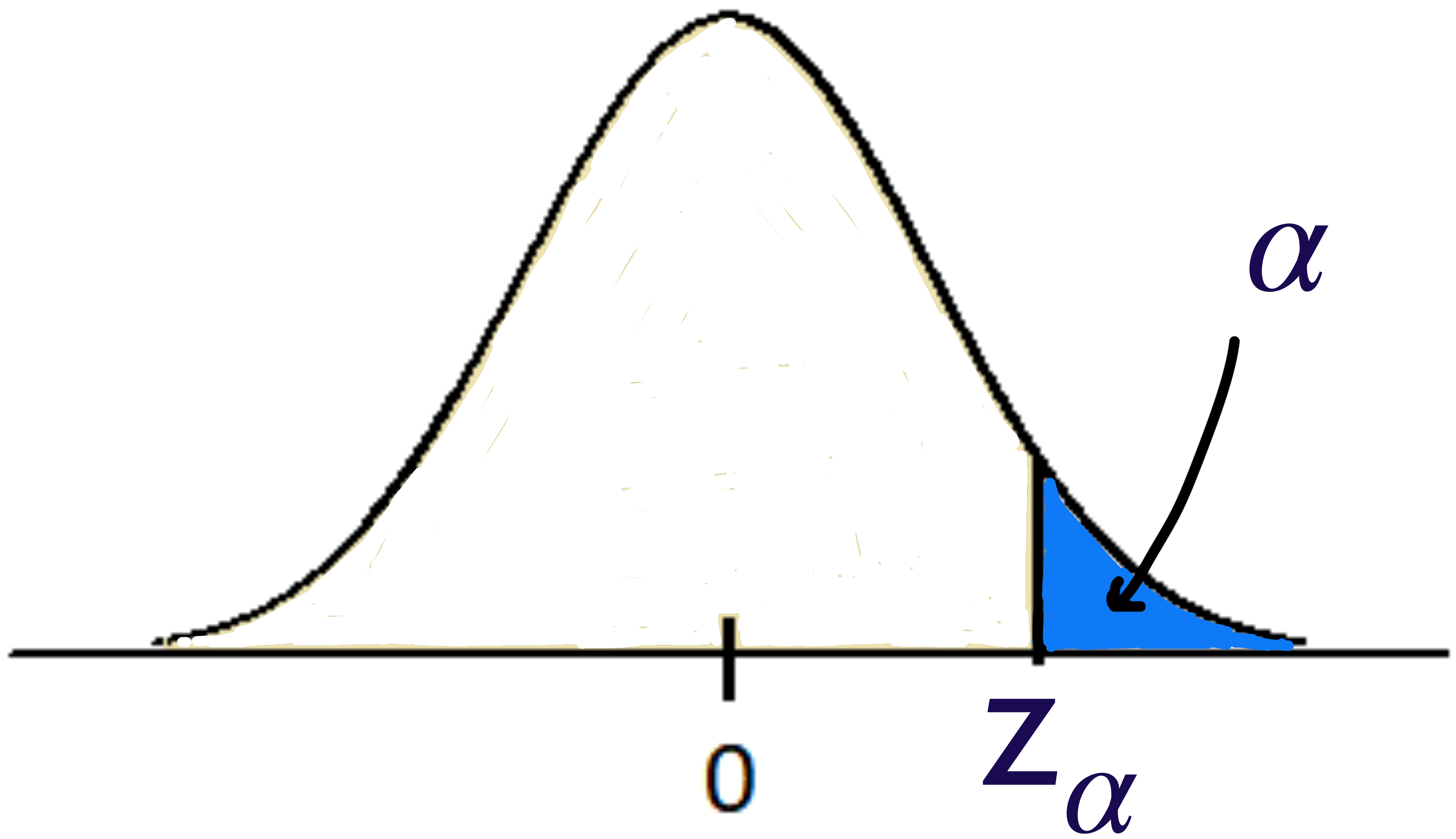
Critical Values

- values that cut off specified areas under pdfs

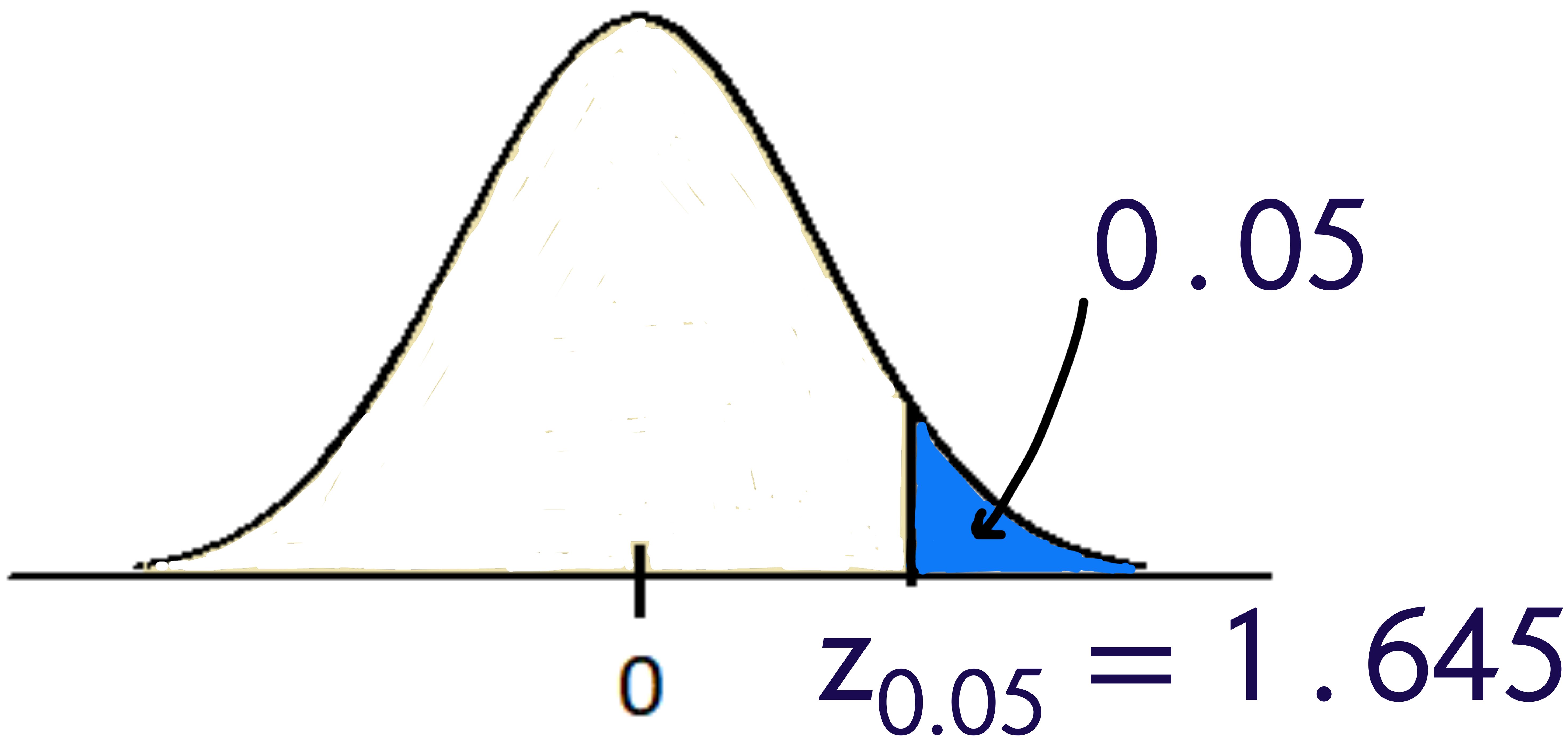
- for the N(0,1) distribution, we will use the notation

$\boxed{z_\alpha}$

to be the value that cuts off area α to the right



Example:



R Code: qnorm(0.95)