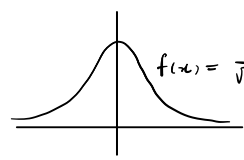
§7.1 Normal distribution

Def A random variable X has the normal distribution with parameters μ and σ^2 if it has density

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for all } x \in \mathbb{R}.$$

Shorthand: $X \sim N(\mu, \sigma^2)$, Mean: $E[X] = \mu$, Variance: $V(X) = \sigma^2$.



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

$X \sim \mathrm{Norm}(0,1)$ density is the famous bell curve

Remark The CDF of X is $F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$

but cannot be given in a closed form expression.

Lemma $\displaystyle\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = 1$

Proof Let $\displaystyle I = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$ $\qquad u = \frac{x-\mu}{\sigma}$

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

$\sigma du = dx$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Observe that

$$I^2 = \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du\right)\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-v^2/2} dv\right)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{(u^2+v^2)}{2}} du\,dv \quad \text{(convert to polar)}$$

$$= \int_{0}^{2\pi}\int_{0}^{\infty} \frac{1}{2\pi} e^{-r^2/2}\, r\,dr\,d\theta$$

$w = -r^2/2$

$dw = -r\,dr$

$$= \frac{-1}{2\pi}\int_0^{2\pi}\left(\int_0^{-\infty} e^w\,dw\right)d\theta$$

$$= \frac{-1}{2\pi}\int_0^{2\pi}\left(e^w\Big|_0^{-\infty}\right)d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} d\theta = 1. \qquad \text{Therefore } I = 1.$$

Lemma If $X \sim N(\mu,\sigma^2)$, then $\frac{X-\mu}{\sigma} \sim N(0,1)$.

Proof Let $Z = \frac{X-\mu}{\sigma}$. We will show that $Z \sim N(0,1)$ by showing it has the correct density. Let $\Phi(x)$ denote the CDF of Z and observe that

$$\begin{aligned}\Phi(x) &= P(Z \le x)\\ &= P\left(\frac{X-\mu}{\sigma} \le x\right)\\ &= P(X \le \mu + \sigma x)\\ &= F_X(\mu+\sigma x).\end{aligned}$$

Therefore the density of Z is given by

$$\begin{aligned}f(x) &= \Phi'(x)\\ &= \frac{d}{dx}\left(F_X(\mu+\sigma x)\right)\\ &= F_X'(\mu+\sigma x)\cdot\sigma \qquad \text{(chain rule)}\\ &= f_X(\mu+\sigma x)\cdot\sigma\\ &= \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(\mu+\sigma x-\mu)^2}{2\sigma^2}}\cdot\sigma\\ &= \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\end{aligned}$$

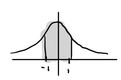
Remark In general, if $X \sim N(\mu,\sigma^2)$, then $aX+b \sim N(a\mu+b,\, a^2\sigma^2)$.

Def $Z \sim N(0,1)$ is called the standard normal random variable. The transformation $\frac{X-\mu}{\sigma}$ when $X \sim N(\mu, \sigma^2)$ is called the standardization of X.

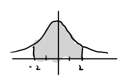
68 - 95 - 99.7 Rule

Through numerical approximation, it is known that when $Z \sim N(0,1)$,

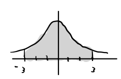
① $P(|Z| \le 1) = P(-1 \le Z \le 1) = 0.68$



② $P(|Z| \le 2) = P(-2 \le Z \le 2) = 0.95$



③ $P(|Z| \le 3) = P(-3 \le Z \le 3) = 0.997$.



Therefore, when $X \sim N(\mu, \sigma^2)$,

① $P(|X-\mu| \le \sigma) = P(|Z| \le 1) \approx 0.68$ "within 1 std dev of the mean"

② $P(|X-\mu| \le 2\sigma) = P(|Z| \le 2) \approx 0.95$ "within 2 std devs of the mean"

③ $P(|X-\mu| \le 3\sigma) = P(|Z| \le 3) \approx 0.997$. "within 3 std devs of the mean"

Example Let $X \sim N(-4, 25)$. Find $P(-9 \le X \le 1)$.

Here $\mu = -4$ and $\sigma = 5$. So

$$P(-9 \le X \le 1) = P\left(\frac{-9-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{1-\mu}{\sigma}\right)$$

$$= P(-1 \le Z \le 1)$$

$$= 0.68$$

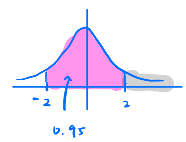
<u>Example</u> Babies' birth weights are normally distributed with mean $\mu = 120$ and standard deviation $\sigma = 20$ ounces. Find the probability a random baby's birth weight is greater than 160 ounces.

Let $X \sim N(120, 20^2)$. Then

$$P(X > 160) = P\left(\frac{X-\mu}{\sigma} > \frac{160-\mu}{\sigma} \right)$$

$$= P(Z > 2)$$

$$\approx \frac{0.05}{2} = 0.025$$



<u>R command</u> $X \sim N(\mu, \sigma^2)$

$P(X \le x) = \text{pnorm}(x, \mu, \sigma)$

Problem 1. Let $X \sim N(-4, 25)$. Find the following probabilities without software.

a. $P(-14 < X < 6)$

b. $P(X > 1)$

c. $P(-9 < X < 6)$

$\mu = -4$, $\sigma = 5$

a. $P\left(\frac{-14-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{6-\mu}{\sigma} \right) = P(-2 < Z < 2) \approx .95$

b. $P\left(\frac{X-\mu}{\sigma} > \frac{1-\mu}{\sigma} \right) = P(Z > 1) \approx \frac{1-0.68}{2} = 0.16$

c. $P\left(\frac{-9-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{6-\mu}{\sigma} \right) = P(-1 < Z < 2) = P(Z < 2) - P(Z \le -1)$

$\approx \left(0.95 + \frac{0.05}{2} \right) - 0.16 = 0.975 - 0.16 = 0.815$

Problem 2. Let $X \sim N(-4, 25)$. Find the following probabilities with R.

a. $P(|X| < 2)$
b. $P(e^X < 1)$
c. $P(X^2 > 3)$

(a) $P(-2 < X < 2) = P(X \le 2) - P(X \le -2)$

(b) $P(X < 0)$

(c) $P(X > \sqrt{3}) + P(X < -\sqrt{3}) = 1 - P(X \le \sqrt{3}) + P(X < -\sqrt{3})$

Problem 2

```{r}
m = -4; s = 5
pnorm(2,m,s) - pnorm(-2,m,s)
pnorm(0,m,s)
1-pnorm(sqrt(3),m,s) + pnorm(-sqrt(3),m,s)
```

```
[1] 0.2295086
[1] 0.7881446
[1] 0.8007507
```

Problem 3. The length of human pregnancy is normally distributed with mean $\mu = 270$ and standard deviation $\sigma = 10$ days. Find the probability that a random pregnancy takes longer than 290 days or less than 240 days.

$P(X > 290) + P(X < 240)$

$= 1 - P(X \le 290) + P(X < 240)$

Problem 3

```{r}
m = 270; s = 10
1-pnorm(290,m,s) + pnorm(240,m,s)
```

```
[1] 0.02410003
```