Module 4: Normal Distribution

Ruben Zamar
Department of Statistics
UBC

January 31, 2016

STANDARD NORMAL

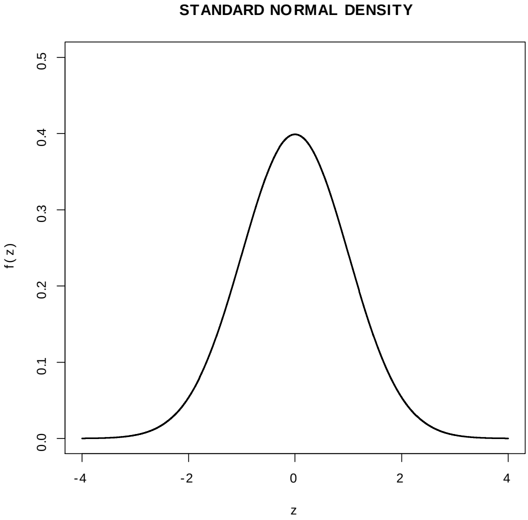
- The standard normal random variable is denoted by Z

- The standard normal density:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-(z^2/2)}, \qquad -\infty < z < \infty$$

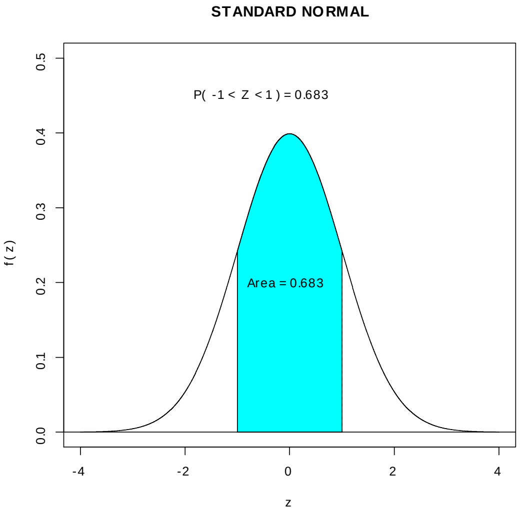
- φ is the "Greek letter" version of f

STANDARD NORMAL DENSITY

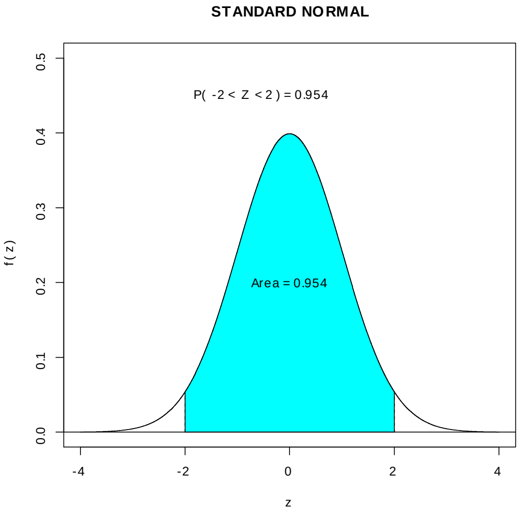


Carl Friedrich Gauss (1777-1855)

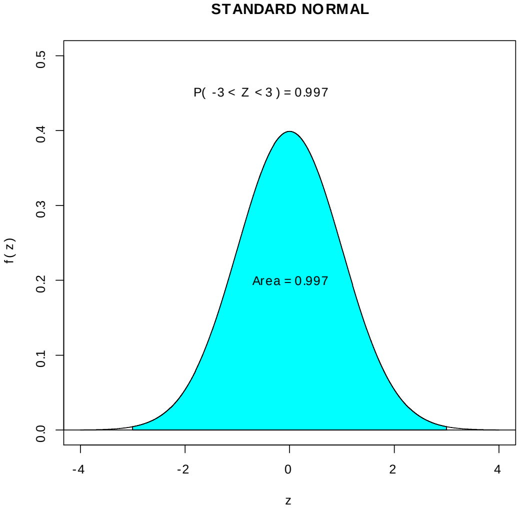
STANDARD NORMAL PROBABILITY



STANDARD NORMAL PROBABILITY



STANDARD NORMAL PROBABILITY



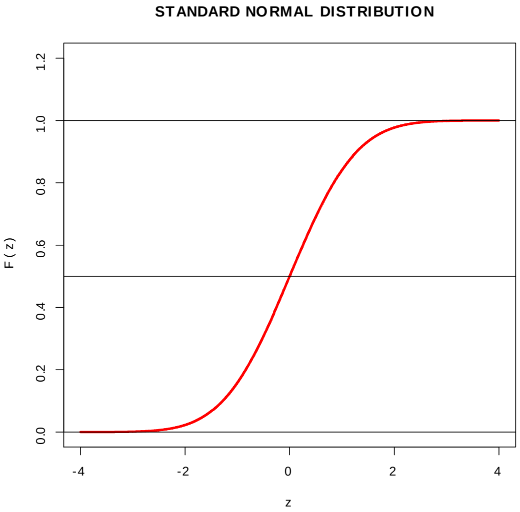
STANDARD NORMAL DISTRIBUTION FUNCTION

$$\Phi(z) = \int_{-\infty}^{z} \varphi(t)dt$$

NOTE 1: denoted by the Greek letter Φ (Greek F)

NOTE 2: $\Phi(z)$ CANNOT BE CALCULATED IN CLOSED FORM

STANDARD NORMAL DISTRIBUTION



STANDARD NORMAL (TABLE AND pnorm)

- "**STANDARD NORMAL TABLE"**: $\Phi(z)$ IS CALCULATED (BY NUMERICAL MEANS) FOR MANY VALUES OF $z > 0$. THE RESULTS ARE DISPLAYED IN THE NORMAL TABLE.

- **BETTER:** USE STATISTICAL SOFTWARE TO EVALUATE $\Phi(z)$. FOR EXAMPLE USE THE FUNCTION **pnorm(z)** IN R

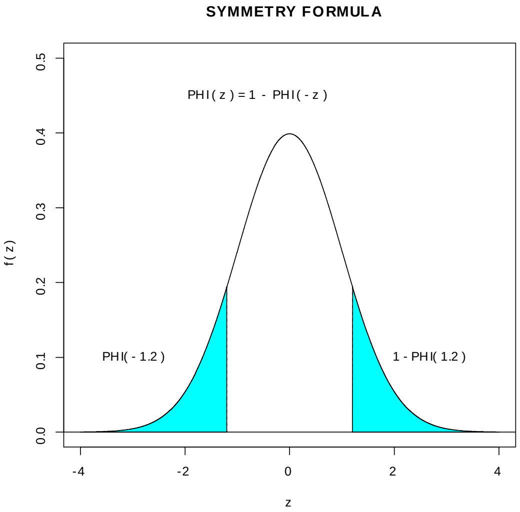
- SYMMETRY FORMULA:

$$\Phi(z) = 1 - \Phi(-z)$$

FOR EXAMPLE

$$\Phi(-1.2) = 1 - \Phi(1.2)$$

ILLUSTRATION OF THE SYMMETRY FORMULA



MEAN, VARIANCE AND STANDARD DEVIATION

$$E(Z) = \int_{-\infty}^{\infty} z\varphi(z)dz = 0$$

$$Var(Z) = \int_{-\infty}^{\infty} z^2\varphi(z)dz = 1$$

$$SD(Z) = 1$$

NOTATION

The notation

$$Z \quad \sim \quad N(0,1) \qquad \text{means}$$

"Z is a normal random variable
with mean 0 and variance 1"

or, equivalently,

"Z is a standard normal random variable"

We have measurements $X_1, X_2, ..., X_n$

$$\boxed{X_i = \mu + \sigma\ Z_i \qquad i = 1, 2, ..., n}$$

MEASUREMENT ERROR MODEL

We have measurements $X_1, X_2, ..., X_n$

$$\boxed{X_i = \mu + \sigma \ Z_i \qquad i = 1, 2, ..., n}$$

X_i	The ith measurement
μ	The "true value"
σ	The "inverse precision" of the measurements
σZ_i	Measurement error in the original scale
Z_i	Measurement error in the "standardized scale"

MEASUREMENT ERROR MODEL

We have measurements $X_1, X_2, ..., X_n$

$$X_i = \mu + \sigma \; Z_i \qquad i = 1, 2, ..., n$$

X_i	The ith measurement
μ	The "true value"
σ	The "inverse precision" of the measurements
σZ_i	Measurement error in the original scale
Z_i	Measurement error in the "standardized scale"

In many applications one assumes that $Z_i \sim N(0, 1)$

STANDARDIZED ERROR

$$Z_i = \frac{X_i - \mu}{\sigma}, \quad i = 1, 2, ..., n$$

FOR INSTANCE

$$Z_i = 1.3$$

MEANS

"THE i^{th} MEASUREMENT IS **1.3** STANDARD UNITS **OVER** THE TRUE VALUE"

$$Z_i \quad = \quad -1.5$$

MEANS

"THE i^{th} MEASUREMENT IS **1.5** STANDARD UNITS **BELOW** THE TRUE VALUE"

GENERAL NORMAL RANDOM VARIABLES

$$X = \mu + \sigma Z \quad , \quad Z \sim N(0,1) \quad \Leftrightarrow \quad Z = \frac{X-\mu}{\sigma}$$

Hence

$$E(X) = E(\mu + \sigma Z) = \mu + \sigma \overbrace{E(Z)}^{0} = \mu$$

$$Var(X) = Var(\mu + \sigma Z) = \sigma^2 \overbrace{Var(Z)}^{1} = \sigma^2$$

NOTATION

Notation:

$$X \sim N\left(\mu, \sigma^2\right) \qquad \text{means}$$

"X is a normal random variable
with mean μ and variance σ^2"

Derive the distribution function for X :

$$
\begin{aligned}
F(x) &= P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\
&= P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)
\end{aligned}
$$

DENSITY

Recall that

$$\Phi'(z) \quad = \quad \varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

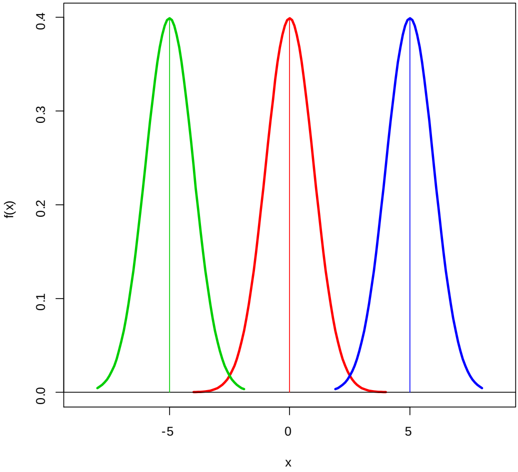
Differentiate

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

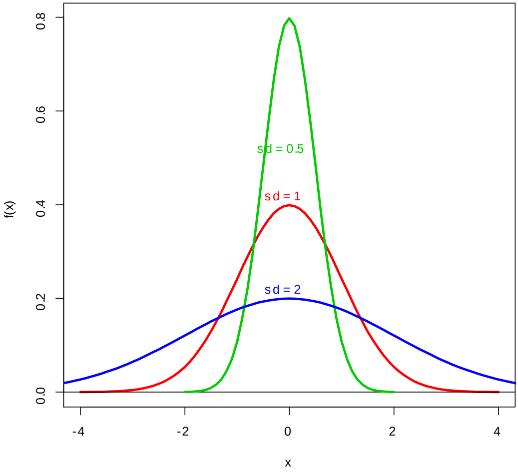
to obtain the density function:

$$f(x) = F'(x) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

DIFFERENT MEANS



DIFFERENT VARIANCES



NORMAL PROBABILITY CALCULATION

THE KEY RESULT TO CARRY ON NORMAL PROB CALCULATIONS IS:

$$X \sim N\left(\mu, \sigma^2\right) \quad \Rightarrow \quad F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

The function $\Phi(z)$ is tabulated and available in R and Matlab.

PRACTICE

Problem 1: Let $Z \sim N(0,1)$. Calculate

a) $P(0.10 \leq Z \leq 0.35)$

b) $P(Z > 1.25)$

c) $P(Z > -1.20)$

d) Find c such that $P(Z > c) = 0.05$

e) Find c such that $P(|Z| < c) = 0.95$

ANSWER TO PROBLEM 1

Part a) $P(0.10 \leq Z \leq 0.35) = ?$

$$\begin{aligned} P(0.10 \leq Z \leq 0.35) &= \Phi(0.35) - \Phi(0.10) \\ &= 0.0970 \end{aligned}$$

In R:

$$\text{pnorm}(0.35) - \text{pnorm}(0.10)$$

ANSWER TO PROBLEM 1 (cont)

Part b) $P(Z > 1.25) = ?$

$$
\begin{aligned}
P(Z > 1.25) &= 1 - P(Z \leq 1.25) \\
&= 1 - \Phi(1.25) = 0.1056
\end{aligned}
$$

In R:

$$1 - \text{pnorm}(1.25)$$

ANSWER TO PROBLEM 1 (continued)

Part c) $P\left(Z > -1.2\right) = ?$

$$
\begin{aligned}
P\left(Z > -1.2\right) &= 1 - P\left(Z \leq -1.2\right) \\
&= 1 - \Phi\left(-1.2\right) \\
&= 1 - \left[1 - \Phi\left(1.2\right)\right] \\
&= \Phi\left(1.2\right) = 0.8849
\end{aligned}
$$

In R:

$$\text{pnorm}\left(1.2\right)$$

PROBLEM 1 (continued)

Part d) Find c such that $P(Z > c) = 0.05$

$$\begin{aligned} 1 - \Phi(c) &= 0.05 \\ \Phi(c) &= 0.95 \\ c &= \Phi^{-1}(0.95) \end{aligned}$$

In R:

$qnorm\,(0.95)$ (inverse for the standard normal cdf function)

ANSWER TO PROBLEM 1 (continued)

Part e) Find c such that $P(|Z| < c) = 0.95$

$$
\begin{aligned}
P(|Z| > c) &= P(-c < Z < c) = \Phi(c) - \Phi(-c) \\
&= \Phi(c) - [1 - \Phi(-c)] = 2\Phi(c) - 1 = 0.95 \\
\Phi(c) &= \frac{1.95}{2} = 0.975 \\
c &= \Phi^{-1}(0.975) \quad \text{(now use the Table)} \\
&= qnorm(0.975) = 1.95996 = 1.96 \quad \text{(using R)}
\end{aligned}
$$

PRACTICE

Problem 2: Let $X \sim N(3, 25)$.

(a) CALCULATE $P(X > 4)$

(b) CALCULATE $P(2 < X < 4)$

(c) CALCULATE $P(X < 1)$

(d) FIND c SUCH THAT $P(X > c) = 0.10$

(e) FIND c SUCH THAT $P(|X - 3| < c) = 0.95$

(f) FIND c SUCH THAT $P(X > c) = 0.90$

RECALL THE KEY EQUATION: IF $X \sim N\left(\mu, \sigma^2\right)$ THEN:

$$F(x) \quad = \quad \Phi\left(\frac{x-\mu}{\sigma}\right)$$

(a) CALCULATE $P(X > 4)$

$$\begin{aligned} P(X &> 4) = 1 - P(X \leq 4) = 1 - F(4) \\ &= 1 - \Phi\left(\frac{4-3}{5}\right) = 1 - \Phi(0.20) = 0.4207 \end{aligned}$$

ANSWER TO PROBLEM 2 (continued)

(b) CALCULATE $P(2 < X < 4)$

$$P(2 < X < 4) = F(4) - F(2) = \Phi\left(\frac{4-3}{5}\right) - \Phi\left(\frac{2-3}{5}\right)$$

$$= \Phi(0.20) - \Phi(-0.20) = 2\Phi(0.20) - 1 = 0.1585$$

(c) CALCULATE $P(X < 1)$

$$P(X < 1) = F(1) = \Phi\left(\frac{1-3}{5}\right) = \Phi(-0.40)$$

$$= 1 - \Phi(0.40) = 0.3446$$

ANSWER TO PROBLEM 2 (continued)

(d) FIND c SUCH THAT $P(X > c) = 0.10$

$$
\begin{aligned}
1 - F(c) &= 0.10 \\
1 - \Phi\left(\frac{c-3}{5}\right) &= 0.10 \\
\Phi\left(\frac{c-3}{5}\right) &= 0.90 \\
\frac{c-3}{5} &= \Phi^{-1}(0.90) = 1.2816 \\
c &= 1.2816 \times 5 + 3 = 9.408
\end{aligned}
$$

(e) FIND c SUCH THAT $P(|X - 3| < c) = 0.95$

(e) FIND c SUCH THAT $P(|X-3|<c)=0.95$

$$
\begin{aligned}
P(-c &< X-3<c)=0.95 \\
P(-c+3 &< X<c+3)=0.95 \\
F(c+3)-F(-c+3) &= 0.95 \\
\Phi\left(\frac{c}{5}\right)-\Phi\left(-\frac{c}{5}\right) &= 0.95 \Rightarrow 2\Phi\left(\frac{c}{5}\right)-1=0.95 \\
\Phi\left(\frac{c}{5}\right) &= \frac{1.95}{2}=0.975 \\
\frac{c}{5} &= \Phi^{-1}(0.975)=1.96 \Rightarrow c=1.96\times 5=9.8
\end{aligned}
$$

ANSWER TO PROBLEM 2 (continued)

(f) FIND c SUCH THAT $P(X > c) = 0.90$

$$
\begin{aligned}
1 - F(c) &= 0.90 \\
1 - \Phi\left(\frac{c-3}{5}\right) &= 0.90 \\
\Phi\left(\frac{c-3}{5}\right) &= 0.10 \\
\frac{c-3}{5} &= \Phi^{-1}(0.10) = -1.2816 \\
c &= -1.2816 \times 5 + 3 = -3.408
\end{aligned}
$$

PRACTICE

Problem 3. A machine fills "10-pound bags" of dry concrete mix. The actual weight of the concrete mix put into the bag is a normal random variable with standard deviation $\sigma = 0.10$ pound. The mean can be set by the machine operator.

(a) What is the mean at which the machine should be set if at most 10% per cent of the bags can be underweight?

(b) What if $\sigma = 0.1\mu$

(a)

- Let X represent the actual weight of the concrete mix in a "10 pound" bag.
- By assumption, $X \sim N(\mu, 0.01)$.
- We should set μ at a value such that

$$P(X < 10) = 0.1$$

ANSWER TO PROBLEM 3 (cont)

$$
\begin{aligned}
P(X < 10) &= 0.1 \\
F(10) &= 0.1 \\
\Phi\left(\frac{10-\mu}{0.1}\right) &= 0.1 \\
\frac{10-\mu}{0.1} &= \overbrace{\Phi^{-1}(0.1) = -1.2816}^{\text{use qnorm}(0.1)\text{ in R}} \\
\mu &= 10 + 0.1 \times 1.2816 = 10.128
\end{aligned}
$$

ANSWER TO PROBLEM 3 (continued)

(b)

- By assumption, $X \sim N\left(\mu, 0.01\mu^2\right)$.

- We should set μ at a value such that

$$P\left(X < 10\right) = 0.1$$

$$
\begin{aligned}
P(X < 10) &= 0.1 \\
F(10) &= 0.1 \\
\Phi\left(\frac{10-\mu}{0.1\mu}\right) &= 0.1 \\
\frac{10-\mu}{0.1\mu} &= \overbrace{\Phi^{-1}(0.1) = -1.2816}^{\text{use qnorm(0.1) in R}} \\
\mu\,(1 - 1.2816 \times 0.1) &= 10 \\
\mu &= \frac{10}{1 - 1.2816 \times 0.1} = 11.470
\end{aligned}
$$