

# The normal distribution, chapter 3.4

Grethe Hystad  
University of Arizona

September 30, 2012

# The Normal distribution

- The normal distribution is the most frequently occurring continuous probability distributions.
- Much statistics theory is based on the normal distribution.
- Central limit theorem

## Examples:

- Heights of adults
- Observations errors in an experiment
- Velocities of the molecules in the ideal gas.
- Many naturally occurring measurements have relative frequency distributions that closely fit the normal curve.

# The Normal distribution

## Theorem

*The random variable  $X$  has a Normal distribution if its p.d.f is defined by*

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

*We have*

$$E(X) = \mu \text{ and } \text{Var}(X) = \sigma^2.$$

*We say that  $X$  is  $N(\mu, \sigma^2)$ .*

# The Normal probability density function

R-code: `plot(function(x) dnorm(x,6,2),-1,13)`

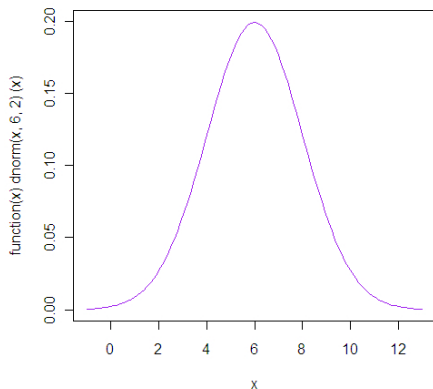


Figure:  $\mu = 6, \sigma = 2$



# The Standard Normal distribution

## Definition

If  $Z$  is  $N(0, 1)$  we say that  $Z$  has a standard normal distribution.

- If  $Z$  is  $N(0, 1)$ , we have

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{w^2}{2}} dw.$$

# The Standard Normal distribution

## Definition

If  $Z$  is  $N(0, 1)$  we say that  $Z$  has a standard normal distribution.

- If  $Z$  is  $N(0, 1)$ , we have

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{w^2}{2}} dw.$$

- Tables of integrals in Appendix, Table Va, gives numerically values for  $\Phi(z) = P(Z \leq z)$  for  $z \geq 0$ .

# The Standard Normal distribution

## Definition

If  $Z$  is  $N(0, 1)$  we say that  $Z$  has a standard normal distribution.

- If  $Z$  is  $N(0, 1)$ , we have

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{w^2}{2}} dw.$$

- Tables of integrals in Appendix, Table Va, gives numerically values for  $\Phi(z) = P(Z \leq z)$  for  $z \geq 0$ .
- For  $z > 0$ , we have  
 $\Phi(-z) = P(Z \leq -z) = P(Z > z) = 1 - \Phi(z)$ .

# The Standard Normal distribution

## Definition

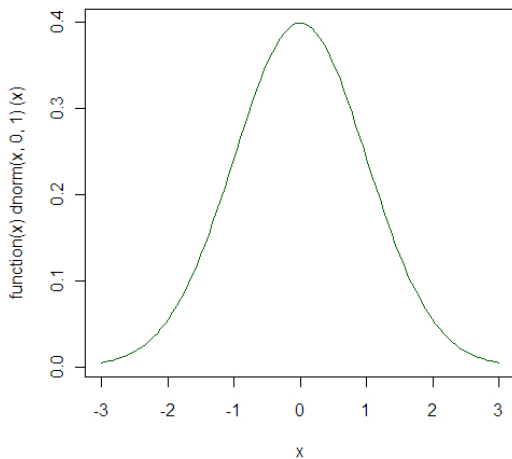
If  $Z$  is  $N(0, 1)$  we say that  $Z$  has a standard normal distribution.

- If  $Z$  is  $N(0, 1)$ , we have

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{w^2}{2}} dw.$$

- Tables of integrals in Appendix, Table Va, gives numerically values for  $\Phi(z) = P(Z \leq z)$  for  $z \geq 0$ .
- For  $z > 0$ , we have  
 $\Phi(-z) = P(Z \leq -z) = P(Z > z) = 1 - \Phi(z)$ .
- Tables of integrals in Appendix, Table Vb, gives numerically values for  $P(Z > z_\alpha) = \alpha$

# The Standard Normal probability density function



# The Standard Normal distribution, example

## Example

Let  $Z$  denote a standard normal variable. Find

(A)  $P(Z \leq 1)$

(B)  $P(Z \geq 1)$

(C)  $P(Z < -1.5)$

(D)  $P(-1 \leq Z \leq 1.5)$

# The Standard Normal distribution, example

## Solution

(A)  $P(Z \leq 1) = \Phi(1) = 0.8413$  (table Va)

(B)  $P(Z \geq 1) = 1 - P(Z \leq 1) = 1 - 0.8413 = 0.1587$  (table Va)  
or  $P(Z \geq 1) = 0.1587$  (table Vb).

(C)  $P(Z < -1.5) = P(Z > 1.5) = 0.0668$  (table Vb)

(D)

$$\begin{aligned}P(-1 \leq Z \leq 1.5) &= \Phi(1.5) - \Phi(-1) \\&= P(Z \leq 1.5) - P(Z \leq -1) \\&= P(Z \leq 1.5) - P(Z \geq 1) \\&= 0.9332 - 0.1587 = 0.7745\end{aligned}$$

# The Standard Normal distribution

We use the notation

- $P(Z \geq z_\alpha) = \alpha$
- $P(Z \leq z_\alpha) = 1 - \alpha$
- $P(Z \leq z_{1-\alpha}) = \alpha$
- In general,  $\alpha$  is the probability to the right of  $z_\alpha$  when  $Z$  is  $N(0, 1)$ .

# The inverse of the Standard Normal distribution, example

## ■ Example

Find a value  $z_0$  such that  $P(Z \leq z_0) = 0.85$ .

R-code:

```
> z0 = qnorm(0.85)  
[1] 1.0364
```

# The inverse of the Standard Normal distribution, example

## Example

Find a value  $z_0$  such that  $P(Z \leq z_0) = 0.85$ .

## Solution

$z_0 = z_{1-\alpha} = z_{0.15} = 1.04$  (table Va)

or we can write  $P(Z \leq z_0) = 0.85$  as  $P(Z \geq z_\alpha) = 0.15$

which again give  $z_0 = z_{0.15} = 1.04$  (table Vb).

R-code:

```
> z0 = qnorm(0.85)
```

```
[1] 1.0364
```

# The Standard Normal distribution, the R-code

- If  $P(Z \leq z_0) = p$ , then in R,  
>  $z_0 = qnorm(p)$ .
- To compute  $F(z_0) = \int_{-\infty}^{z_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw$ , in R:  
>  $pnorm(z_0)$ .
- Example: To find  $P(Z \leq 1)$ , we compute  
>  $pnorm(1)$   
[1] 0.8413.

# The Standard Normal distribution

We have

$$P(-1 \leq Z \leq 1) = P(Z \leq 1) - P(Z \leq -1) = 0.8413 - 0.1587 = 0.683.$$

$$P(-2 \leq Z \leq 2) = 0.954.$$

$$P(-3 \leq Z \leq 3) = 0.997.$$

Thus, for a standard normal distribution, approximately, 68% of the values fall within 1 standard deviation of the mean in either direction.

95% of the values fall within 2 standard deviation of the mean in either direction.

99.7% of the values fall within 3 standard deviation of the mean in either direction.

# The Standard Normal distribution

Define the z-score (standardized variable)

$$Z = \frac{X - \mu}{\sigma},$$

where  $X$  is a random variable that have a normal distribution with mean  $\mu$  and standard deviation  $\sigma$ .

## Theorem

*If  $X$  is  $N(\mu, \sigma^2)$ , then*

$$Z = \frac{X - \mu}{\sigma}$$

*is  $N(0, 1)$*

Proof given in class.

# The Normal distribution

Suppose  $X$  is  $N(\mu, \sigma^2)$ . Then

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

where  $Z = \frac{X - \mu}{\sigma}$  is  $N(0, 1)$ .

# The Normal distribution

## Example

Find  $P(4 \leq X \leq 12)$ , where  $X$  is  $N(8, 4)$ .

## Solution

Let

$$Z = \frac{X - 8}{2},$$

where  $\mu = 8$  and  $\sigma = 2$ . Then

$$\begin{aligned} P(4 \leq X \leq 12) &= P\left(\frac{4 - 8}{2} \leq Z \leq \frac{12 - 8}{2}\right) \\ &= P(-2 \leq Z \leq 2) = 0.954. \end{aligned}$$

# The Normal distribution, R-code

In R, the problem can also be solved in this way:

$$P(4 \leq X \leq 12) = P(X \leq 12) - P(X \leq 4).$$

Then type

```
> pnorm(12,8,2)-pnorm(4,8,2)
[1] 0.954.
```

In general, in R:

```
> pnorm(x, μ, σ)
```

which is the cumulative normal distribution,  $P(X \leq x)$ , with mean  $\mu$  and standard deviation  $\sigma$ .

# The Normal distribution

## Example

Find  $P(|X - 8| \leq 2)$ , where  $X$  is  $N(8, 4)$ .

## Solution

$$\begin{aligned}P(|X - 8| \leq 2) &= P(-2 < X - 8 < 2) \\&= P\left(-1 < \frac{X - 8}{2} < 1\right) \\&= P(-1 < Z < 1) \\&= 0.683.\end{aligned}$$

# The Chi square distribution

## Theorem

If the random variable  $X$  is  $N(\mu, \sigma^2)$  and  $\sigma^2 > 0$ , then the random variable  $V = \left(\frac{X-\mu}{\sigma}\right)^2 = Z^2$  is  $\chi^2(1)$ .

## Example

Suppose  $Z$  is  $N(0, 1)$ . Then  
 $P(|Z| < \sqrt{0.016}) = P(Z^2 < 0.016) = 0.100$ , using the chi-square table with  $r = 1$ .

# The Normal distribution

## Theorem

*If  $X$  has a normal distribution with mean  $\mu$  and a variance of  $\sigma^2$ , and*

$$Y = aX + b,$$

*where  $a, b$  are constants, then  $Y$  also has a normal distribution with*

$$E(Y) = a\mu + b \text{ and } \text{Var}(Y) = a^2\sigma^2.$$

Proof left as a homework problem.

# The Normal distribution, inflection points

- The probability density function,  $f(x)$ , of a normal random variable with a mean of  $\mu$  and a standard deviation of  $\sigma$  has points of inflection at  $x = \mu \pm \sigma$ . The proof of this is left as a homework problem.

## Example

If  $X$  has a normal distribution with mean 2 and standard deviation 2, find  $P(X^2 - 4X \leq 12)$ .

## Solution

$$\begin{aligned}P(X^2 - 4X \leq 12) &= P(X^2 - 4X + 4 \leq 16) \\&= P((X - 2)^2 \leq 16) \\&= P(-4 \leq (X - 2) \leq 4) \\&= P(-2 \leq \frac{X - 2}{2} \leq 2) \\&= P(-2 \leq Z \leq 2) = 0.95,\end{aligned}$$

where  $Z = \frac{X-2}{2}$ .



# The Normal distribution, example

## Example

The math SAT 1 scores among U.S. college students is approximately normally distributed with a mean of 500 and standard deviation of 100. What is the probability that the SAT score for a randomly selected student is greater than 550?

# The Normal distribution, example

## Example

The math SAT 1 scores among U.S. college students is approximately normally distributed with a mean of 500 and standard deviation of 100. What is the probability that the SAT score for a randomly selected student is greater than 550?

## Solution

Let  $X$  be the SAT score of the student. We have

$$P(X > 550) = P\left(Z > \frac{550-500}{100}\right) = P(Z > 0.5) = 0.3085, \text{ where}$$

$$Z = \frac{X-500}{100}.$$

# The Normal distribution, example

## Problem

*Suppose a university requires a math SAT score of at least 550 for admission. 90% of the students that apply are from high schools with a mean math SAT score of 500 and standard deviation of 100. 10% of the students are from a high school, called high school A, in which the mean math SAT score is 600 with a standard deviation of 50. The SAT scores among both populations are approximately normally distributed. Suppose a randomly selected applicant is selected. He/she meets the SAT requirement. What is the probability that the applicant is from high school A?*

Solution given in class.