

# Set 15 - The Normal Distribution

March 1, 2023 12:20 PM

## Stat 260 Lecture Notes Sets 15 - The Normal Distribution

In this section we will look at a specific **continuous distribution** called the **normal distribution**. The pdf for the normal distribution is

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

try graphing  $\mu=0, \sigma=1$   
 $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$   
 all just numbers

where  $-\infty < x < \infty$ , and  $\mu$  = the mean, and  $\sigma$  = the standard deviation.

As we saw in Sets 13 and 14, we will be using the cdf to calculate probabilities for a continuous random variable (not the pdf), so we will not need this  $f(x)$  function. Instead we will only use a **cdf table** to calculate our probabilities here.

don't have to integrate

How to picture the normal distribution:



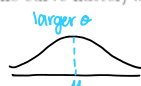
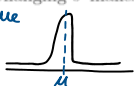
- the **bell curve**
- symmetric about  $\mu$
- **inflection points** at one standard deviation away from the mean (at  $x = \mu \pm \sigma$ )

The exact shape of the curve depends on the values of  $\mu$  and  $\sigma$ .

Changing  $\mu$  moves the picture left or right. (changes where peak is)

Changing  $\sigma$  makes the curve flatter/wider or taller/skinnier. (changes where inflection points are)

small value of  $\sigma$



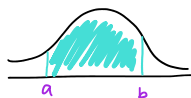
larger  $\sigma$

area under curve has to be 1 for both

Note: The normal distribution is continuous so **all our rules from Set 13 and 14 apply** (the area under the curve is 1,  $f(x) \geq 0$  for all  $x$ ,  $P(a \leq X \leq b)$

is the area under the  $f(x)$  curve from  $x = a$  to  $x = b$ ).

$$P(a \leq X \leq b) = \text{area}$$



online assignment 2 (can do most questions already)  
 Test 2: sets 10 - 17  
 need calculator  
 given formula sheet and distribution tables  
 8 short answer, 1 long answer  
 practice midterms

For a continuous random variable  $X$ , to calculate  $P(X \leq x)$  by hand using the pdf  $f(x)$  we would need calculus. Notice that

$$F(x) = \int_{-\infty}^x f(y) dy = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy$$

we never use this

is a very difficult integral - we would need numerical methods to evaluate this.

Instead we will use our cdf table to find  $P(X \leq x)$ . (In other words, you will never have to write down the  $f(x)$  function to solve a question with the normal distribution!)

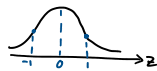
When we had cdf tables for the binomial distribution we had many tables for different combinations of the values of  $n$  and  $p$ . We could do the same for the normal distribution but then we would need many, many, many tables for all the different combinations of values of  $\mu$  and  $\sigma$  (remember,  $\mu$  and  $\sigma$  do not have to be whole numbers!). This would result in needing a massive book of tables to solve our questions.

Instead, we "standardize" the particular distribution we are working with and then we only need one table.

If  $Z$  is a normal random variable with  $\mu = 0$  and  $\sigma = 1$  then it is called a **standard normal random variable**. (Here we use  $Z$  instead of  $X$  to highlight the fact that the variable has been standardized. We will use  $X$  for our original distribution and  $Z$  for the standardized distribution.)

We have a cdf table to calculate  $P(Z \leq z)$  for a standard normal random variable  $Z$ .

$Z$  :



transform question into standardized then use the tables

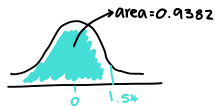
→ Shift & Scale

→  $Z$  indicates it's the standard form

**Example 1:** Find the following probabilities for the standard normal random variable  $Z$ .

(a)  $P(Z \leq 1.54) = 0.9382$

$\downarrow 1.5 \rightarrow 0.04$



$$P(Z \geq 1.54) = 1 - P(Z \leq 1.54) = 1 - 0.9382$$



(b)  $P(Z \leq -1.69) = 0.0455$

How do we find  $P(X \leq x)$  for a normally distributed random variable  $X$  with  $\mu \neq 0$  or  $\sigma \neq 1$ ?

We **standardize** by substituting

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

moves  
scales

**NOT** on formula sheet

-3.49 is smallest value

↳ if value off table small then 0

↳ off table big then 1

**Example 2:** Suppose  $X$  is a normal random variable with  $\mu = 3$  and  $\sigma = 2$ .

(a) Find  $P(X \leq 5)$ .

$$P(X \leq 5) = P\left(\frac{X - \mu}{\sigma} \leq \frac{5 - 3}{2}\right) = P(Z \leq 1) = P(Z \leq 1.00) = 0.8413$$

→ round to 2 decimal places

(b) Find  $P(X < 5)$ . =  $P(X \leq 5) = 0.8413$

**Example 3:** Suppose  $X$  is a normal random variable with  $\mu = 15$  and  $\sigma = 7$ . \*information you need to be given

(a) Find  $P(X \geq 25)$ . → can do complement/standardizing in any order.

$$= P\left(\frac{X - \mu}{\sigma} \geq \frac{25 - 15}{7}\right) = P(Z \geq 1.43) = 1 - P(Z \leq 1.43)$$

round to 2 decimal places

=  $1 - 0.9236$

=  $0.0764$

can have  $\geq$  because it's a continuous random variable

→ b/c tables are only accurate to 2 decimal places

(b) Find  $P(6 \leq X \leq 27)$ . =  $P\left(\frac{6 - 15}{7} \leq \frac{X - \mu}{\sigma} \leq \frac{27 - 15}{7}\right)$

=  $P(-1.29 \leq Z \leq 1.71)$

=  $P(Z \leq 1.71) - P(Z \leq -1.29)$

=  $0.9564 - 0.0985$

=  $0.8579$

rounding negatives

We can use the Z table (the cdf table for the normal distribution) in reverse.

**Example 4:** What is the value of  $z$  so that  $P(Z \leq z) = 0.9900$ ?



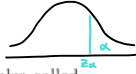
$$P(Z \leq z) = 0.9900$$

$$P(Z \leq 2.33) = 0.9901$$

$$P(Z \leq z) = 0.9900$$

**Notation:**  $z_\alpha$  is the  $z$  value so that the area to the right is  $\alpha$ .

That is,  $z_\alpha$  is the value so that  $P(Z \geq z_\alpha) = \alpha$ . The  $z_\alpha$  value is also called a **critical value**.



**Example 5:** Find  $z_{0.01}$ .

$$P(Z \geq z_{0.01}) = 0.01$$

$$P(Z \leq z_{0.01}) = 0.9900$$

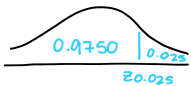
In example 4 we saw that

$$P(Z \leq 2.33) = 0.9901$$

$$\text{So } z_{0.01} = 2.33$$



**Example 6:** Find  $z_{0.025}$ .



look up prob. 0.9750

$$z_{0.025} = 1.96$$

\* **General rule:** pick probability closest to ours on the table, if our value is exactly between two choices then we pick the halfway point.

→ almost never get exact value you're looking for  
→ find closest value on table  
→ find probability closest

Example 7: Find  $z_{0.05}$ .



look up the prob. 0.9500  
 Table has: 0.9495 and 0.9505  
 1.64      1.65  
             1.645  
 on table:  $P(Z \leq 1.64) = 0.9495$   
 $P(Z \leq 1.65) = 0.9505$   
 So  $z_{0.05} = 1.645$ .

*exactly halfway between*

Another way to find  $z_\alpha$ :

Example 8: Find  $z_{0.02}$ .



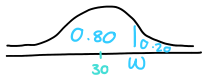
look up  $z$  value for the prob. 0.02.  
 on table  $P(Z \leq -2.05) = 0.02$

using symmetry, we have  
 $P(Z \geq 2.05) = 0.02$   
 So  $z_{0.02} = +2.05$

Notation: "is distributed as"

- $X \sim N(\mu, \sigma)$
- $X \sim N(\mu, \sigma^2)$
- $X \sim \text{binomial}(n, p)$
- $X \sim \text{Poisson}(\lambda)$

**Example 9:** Suppose  $X$  is a normal random variable with mean 30 and standard deviation 17. Find the cutoff value  $w$  so that 20% of the distribution is above  $w$ .



look up the prob. 0.8000 on stat table.

$$Z = 0.84$$

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

$$0.84 = \frac{w - 30}{17}$$

$$w = 17(0.84) + 30 \\ = 44.28$$