The Standard Normal Distribution

Section 4.3

Cathy Poliak, Ph.D.
cathy@math.uh.edu
Office hours: T Th 2:30 - 5:15 pm 620 PGH

Department of Mathematics
University of Houston

February 23, 2016

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Outline

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper Set Up

- Fill in all of the proper bubbles.
- Use a #2 pencil.
- This is popper number 07.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper Questions

An orange juice producer buys all his oranges from a large orange grove. The amount of juice squeezed from each of these oranges is approximately normally distributed, with a mean of 4.70 ounces and a standard deviation of 0.40 ounce. Hint: apply the Empirical Rule (68-95-99.7 rule).

1. What is the probability that an orange from this orange grove has between 3.9 and 5.5 ounces of juice?

   a. 0.68 b. 0.95 c. 0.997 d. 1

2. What is the probability that an orange from this orange grove has less than 4.7 ounces of juice?

   a. 0.68 b. 0.95 c. 0.997 d. 0.5

3. Approximately 95% of the oranges have juice between what two middle values?

   a. 4.3 and 5.1 b. 3.9 and 5.5 c. 3.5 and 5.9 d. 0 and 4.7

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Not 1, 2 or 3 Standard Deviations

An orange juice producer buys all his oranges from a large orange grove. The amount of juice squeezed from each of these oranges is approximately normally distributed, with a mean of 4.70 ounces and a standard deviation of 0.40 ounce. The random variable is X = the amount of juice squeezed from one orange.

- What is the probability that an orange will have less than 4 ounces of juice? $P(X < 4)$.

- There are a couple of ways to answer this question.
  - R: pnorm(x,mean,sd), pnorm(4,4.7,0.4) = 0.04005916
  - TI 83 or 84: normcdf(lowest limit, upper limit, mean, sd), normcdf(-1e99,4,4.7,0.4) = 0.0400591
  - Table A: In the appendix of your book. This table is for Standard Normal Distribution.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Standard Normal Distribution

- If X is an observation from a distribution that has mean μ and standard deviation σ,the standardized value of x is

$$z = \frac{X - \mu}{\sigma} = \frac{\text{observation} - \text{mean}}{\text{standard deviation}}$$

  This is called a $\boldsymbol{z - score}$.
- The standard Normal distribution is the Normal distribution with mean 0 and standard deviation 1 for any variable.

Key concepts for z-scores

- The z-score is the number of standard deviations a value is from the mean.
- Z-scores have no units
- They measure the distance an observation is from the mean in standard deviations.
- Positive z-scores indicate that the observation is above the mean.
- Negative z-scores indicate that the observation is below the mean.
- Z-scores usually are between -3 and 3. Anything beyond these two values indicates that the observation is extreme.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Normal Distribution Calculations

- Area under a Normal curve represent proportions (probability) of observations within a range of values. There is no easy formula to find the area under a Normal curve.

- We use a table or software that calculates the desired areas. The table we use is Table A. It uses a **cumulative proportion**. A cumulative proportion is the proportion (probability) of observations in a distribution that lie at or below a given value.

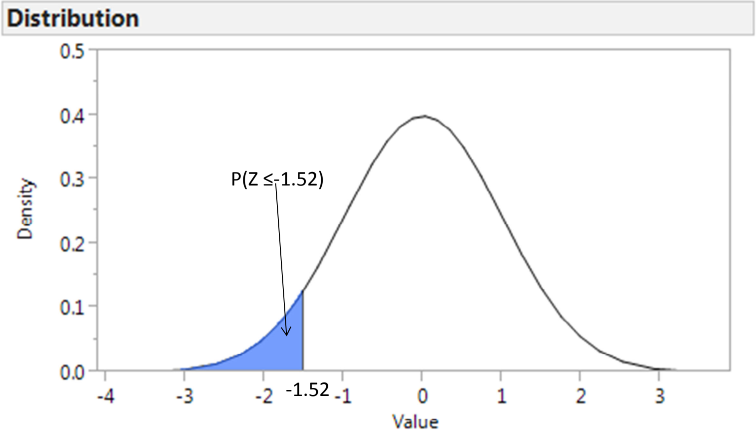
- When the distribution is given by a density curve, the cumulative proportion is the area under the curve to the left of a given value.

Using table A

- The vertical margin are the left most digits of a z-score.
- The top margin is the hundredths place of a z-score.
- The numbers inside the table represents the area from $-\infty$ to that z-score.
- Remember that the standard Normal density curve is symmetric and the total area is equal to 1.
- *Note*: R can calculate these probabilities and also some calculators. Without having to convert to z-scores.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

$P(Z \leq -1.52)$


Distribution
0.5
0.4
0.3
0.2
0.1
0.0
Density
P(Z ≤-1.52)
-4
-3
-2
-1.52
-1
0
1
2
3
Value

$P(Z \leq -1.52) = 0.0643$

R: pnorm(-1.52) = 0.06425549,
TI-83(84):normalcdf(-1e99,-1.52)=0.0642555

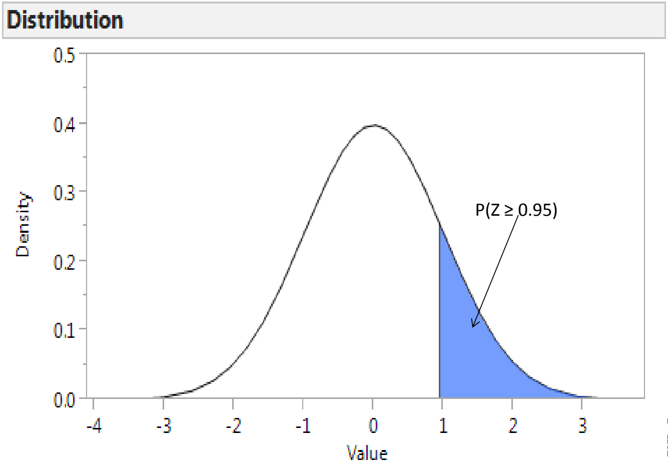
Table A: P(Z < z)

z	0.00	0.01	0.02	0.03
-3.4	0.0003	0.0003	0.0003	0.0003
-3.3	0.0005	0.0005	0.0005	0.0004
-3.2	0.0007	0.0007	0.0006	0.0006
⋮	⋮	⋮	⋮	⋮
-1.5	0.0668	0.0655	0.0643	0.0630

P(Z ≤ -1.52)

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

$P(Z \geq 0.95)$



$P(Z \geq 0.95) = 0.1711$

R: 1 - pnorm(0.95) = 0.1710561,
TI-83(84):normalcdf(0.95,-1e99)=0.1710561

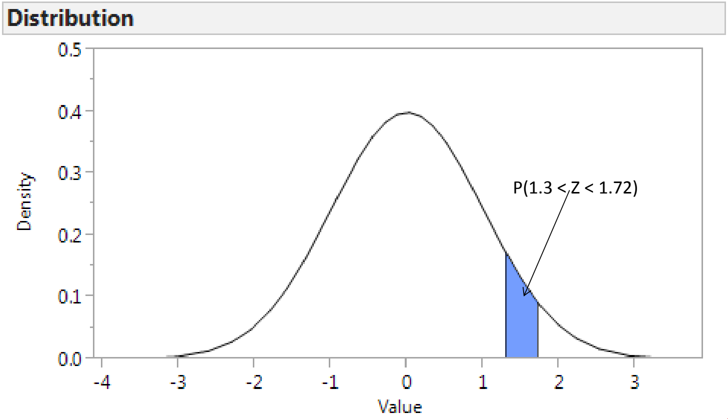
Table A: P(Z < z)

z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289

P(Z ≥ 0.95)= 1 – P(Z < 0.95) = 1 – 0.8289 = 0.1711

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

$P(1.3 < Z < 1.72)$


Distribution
0.5
0.4
0.3
0.2
0.1
0.0
Density
-4
-3
-2
-1
0
1
2
3
Value
P(1.3 < Z < 1.72)

$P(1.3 < Z < 1.72) = 0.0541$

R: pnorm(1.72) - pnorm(1.3) = 0.05408426,
TI-83(84):normalcdf(1.3,1.72)=0.0540843

z	0.00	0.01	0.02	0.03
0.0	0.5000	0.5040	0.5080	0.5120
	⋮	⋮	⋮	⋮
0.1				
1.3	0.9032	0.9049	0.9066	0.9082
	⋮	⋮	⋮	⋮
1.4				
1.7	0.9554	0.9564	0.9573	0.9582

$P(1.3 < Z < 1.72) = 0.9573 - 0.9032 = 0.0541$

Finding Standard Normal probabilities

Using Table A

- The numbers **inside** the table, the four digit numbers between 0 and 1, are the cumulative probabilities or area to the left of a z-score under a standard Normal density curve.
- To determine the probability less than a z-score use the value directly from the table.

$$P(Z < z) = \text{value from table}$$

- To determine the probability for greater than a z-score take one minus the value directly from the table.

$$P(Z > z) = 1 - P(Z < z)$$

- To determine the probability between two values find the difference between the areas directly from the table corresponding to each value.

$$P(z_1 < Z < z_2) = P(Z < z_2) - P(Z < z_1)$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper Questions

Find the following probabilities using Table A, R or your calculator.

4. $P(Z \leq -0.92)$

   a. 0.92 b. 0.1788 c. 0.8212 d. -0.1788

5. $P(Z \leq 1.35)$

   a. 0.135 b. 0.9115 c. 0.0885 d. 0

6. $P(Z \geq 1.96)$

   a. 0.975 b. 0.025 c. -0.975 d. -0.025

7. $P(-0.92 \leq Z \leq 1.96)$

   a. 0.1788 b. 0.975 c. 0.7962 d. -0.7962

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Finding probabilities for Normal distribution

1. State the problem in terms of a probability statement $P(X < x)$, $P(X > x)$, $P(x_1 < X < x_2)$.
2. Standardize X to state the problem in terms of a z-score. (If using the table)
$$z = \frac{x - \mu}{\sigma}$$
3. Draw a picture to show the area under the standard Normal curve.
4. Find the required area under the standard Normal curve using table A (in front of your book).

Probability of amount of juice

The amount of juice squeezed from each of these oranges in an orange grove is approximately Normally distributed, with a mean of 4.70 ounces and a standard deviation of 0.40 ounces. What is the probability that an orange squeezed from this grove has more than 5 ounces of juice?

1. State the problem in terms of a probability.

$$P(X > 5)$$

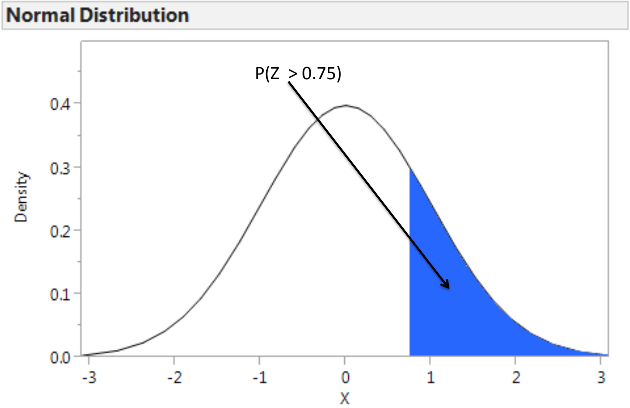
2. Standardize the value.

$$P\left(\frac{(X-\mu)}{\sigma} > \frac{(5-4.7)}{0.4}\right) = P(Z > 0.75)$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

$P(X > 5) = P(Z > 0.75)$

3. Draw a picture to show the desired area under the standard Normal curve.



$P(X > 5) = P(Z > 0.75)$

4. Find the required area under the standard Normal curve using Table A, in R: 1 - pnorm(5,4.7,.4) = 0.2266274 or in TI-83(84):normalcdf(5,1e99,4.7,.4)=0.226627

z	0.00	0.01	...	0.05
0.0	0.5000	0.5040	...	0.5199
0.1	0.5398	0.5438	...	0.5596
0.2	0.5793	0.5832	...	0.5987
0.3	0.6179	0.6217	...	0.6368
0.4	0.6554	0.6591	...	0.6736
0.5	0.6915	0.6950	...	0.7088
0.6	0.7257	0.7291	...	0.7422
0.7	0.7580	0.7611	...	0.7734

P(Z > 0.75) = 1 – P(Z < 0.75) = 1 – 0.7734 = 0.2266

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Probability of amount of juice

The amount of juice squeezed from each of these oranges in an orange grove is approximately Normally distributed, with a mean of 4.70 ounces and a standard deviation of 0.40 ounces. What is the probability that an orange squeezed from this grove has between 3.7 and 4.2 ounces of juice?

1. State the problem in terms of a probability.

$$P(3.7 < X < 4.2)$$

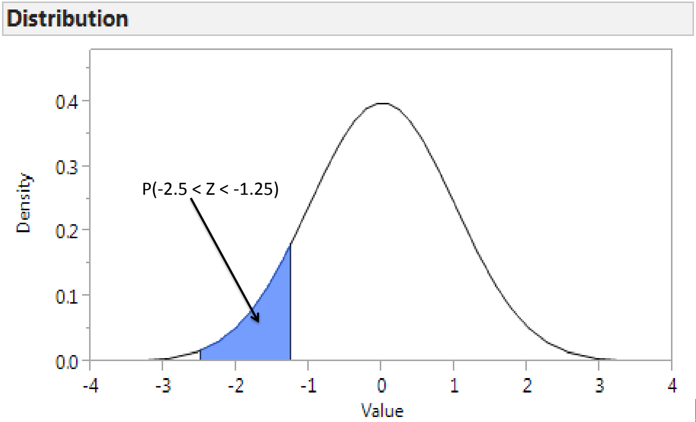
2. Standardize the values.

$$P\left(\frac{(3.7 - 4.7)}{0.4} < \frac{(X - \mu)}{\sigma} < \frac{(4.2 - 4.7)}{0.4}\right) = P(-2.5 < Z < -1.25)$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

$P(3.7 < X < 4.2) = P(-2.5 < Z < -1.25)$

3. Draw a picture to show the desired area under the standard Normal curve.



HOUSTON
ATHEMATICS

$P(3.7 < X < 4.2) = P(-2.5 < Z < -1.25)$

4. Find the required area under the standard Normal curve using Table A , in R: pnorm(4.2,4.7,.4)-pnorm(3.7,4.7,.4) = 0.09944011 or in TI-83(84):normalcdf(3.7,4.2,4.7,.4) = 0.09944.

z	0.00	0.01	0.05
-3.4	0.0003	0.0003	0.0003
⋮	⋮	⋮	⋮
-2.5	0.0062	0.0060	0.0054
⋮	⋮	⋮	⋮
-1.2	0.1151	0.1131	0.1056

P(-2.5 < Z < -1.25) = 0.1056 – 0.0062 = 0.0994

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper Questions

The MPG of a Toyota Prius has a Normal distribution with mean of 49 mpg, $\mu = 49$ and standard deviation 3.5 mpg, $\sigma = 3.5$. Determine the following probabilities using Table A.

8. What is the probability that a Prius has mpg greater than 50 mpg?

   a. 0.6125 b. 0.3875 c. 0.1094 d. 0.6074

9. What is the probability that a Prius has mpg between 40 and 50 mpg?

   a. 0.0051 b. 0.9949 c. 0.3824 d. 0.6074

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Finding a value when given a proportion

- Called inverse Normal.
- This is working “Backwards” using Z-Table.
- Finding the observed values when given a percent.
- In R: qnorm(proportion,mean,sd).
- In TI-83 or 84: invNorm(proportion,mean,sd).

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

“Backward” Normal calculations Using Z-Table

1. State the problem. Since, Z-Table, qnorm and invNorm gives the areas to the left of z-scores, always state the problem in terms of the area to the left of x. Keep in mind that the total area under the standard Normal curve is 1.

2. Use Table A to find c. This is the value from the table not a value that we calculate.

3. Unstandardized to transform the solution from the z-score back to the original x scale. Solving for x using the equation

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

gives the equation $x = \sigma(c) + \mu$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Examples to Work "Backwards" with the Normal Distribution

Find the value of c so that:

1. $P(Z < c) = 0.7704$

2. $P(Z > c) = 0.006$

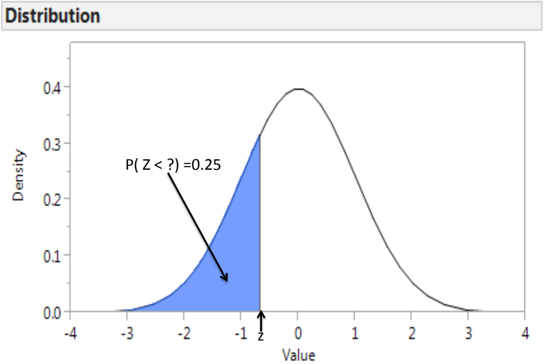
3. $P(-c < Z < c) = 0.966$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

MPG for Prius

The miles per gallon for a Toyota Prius has a Normal distribution with mean $\mu = 49$ mpg and standard deviation $\sigma = 3.5$ mpg. 25% of the Prius have a MPG of what value and lower?

1. We want c, such that $P(Z < c) = 0.25$. That is we want to know what z-score cuts off the lowest 25%.



Find c such that $P(Z < c) = 0.25$

3. From Table A, find something close to 0.25 **inside** the table.

z	0.00	0.01	0.02	0.07	0.08	0.09
-3.4	0.0003	0.0003	...	0.0003	0.0003	0.0002
⋮	⋮	⋮	⋮	⋮	⋮	⋮
-0.7	0.2420	0.2389	...	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	...	0.2514	0.2483	0.2451

P(Z < ?) = 0.25 (closes value is 0.2514)

z = -0.67 (-0.6 “row” + 0.07 “column”)

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Find c such that $P(Z < c) = 0.25$

4. Unstandardized: $x = \sigma(c) + \mu = 3.5(-0.67) + 49 = 46.655$

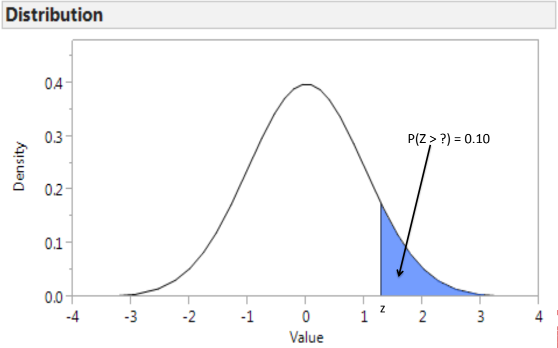
5. This means that 25% of the Prius has a mpg of less than 46.655 mpg.

Using R: qnorm(0.25,49,3.5) = 46.63929,
TI-83(84):invNorm(0.25,49,3.5)=46.63929

Top 10%

Suppose you rank in the 10% of your class. If the mean GPA is 2.7 and the standard deviation is 0.59, what is your GPA? (Assume a Normal distribution)

1. We want c, such that $P(Z > c) = 0.10$. That is we want to know what z-score cuts off the highest 10%.



Find c such that $P(Z > c) = 0.1$

3. From Table A, the areas are **below** or to the left of a z-score thus we want to find something close to 0.90 **inside** the table.

z	0.00	0.01	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5675	0.5714	0.5753
		⋮	⋮	⋮	⋮
0.2	0.5793				
1.2	0.8849	0.8869	0.8980	0.8997	0.9015

P(Z < ?) = 0.90 (close value is 0.8997)

z = 1.28 (1.2 “row” + 0.08 “column”)

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Find c such that $P(Z > c) = 0.1$

4. Unstandardized: $x = \sigma(c) + \mu = 0.59(1.28) + 2.7 = 3.4552$

5. This means that your gpa is 3.4375 if you rank at the 10% of your class.

In R: qnorm(0.9,2.7,0.59) = 3.456115,
TI-83(84):invNorm(0.9,2.7,0.59)=3.456115

Popper Questions

Replacement times for televisions are normally distributed with a mean of 8.2 years and a standard deviation of 1.1 years (based on data from *Getting Things Fixed*, Consumer Reports).

10. If you want to provide a warranty so that only 1% of the televisions will be replaced before the warranty expires, what is the time length of the warranty? These are in years.

a. 6.79 b. 5.64 c. -2.33 d. 8.211

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS