Normal Table Gymnastics

Dr Tom Ilvento
Department of Food and Resource Economics
University of Delaware

Overview

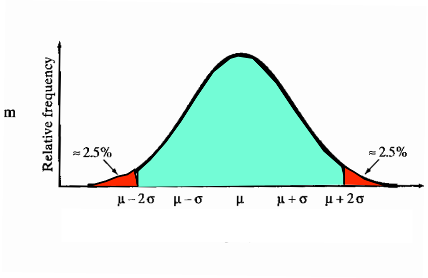
- Let's continue working with the normal table
- And I will show you how to do some table gymnastics to solve for:
  - probabilities out in the tails
  - probabilities between two values
  - calculating the value of X at a certain percentile

Problem

- Suppose a variable is distributed normally with a mean = 300 and a standard deviation of 30
  - X ~ N μ= 300 σ= 30
- What is the probability that a value of x is more than 2 standard deviations away from the mean?
- STEPS:
  - Draw it out
  - Calculate z-score
  - Check the table
  - Do any final calculations

The probability that a value of x is more than 2 standard deviations away from the mean

- **Draw it out**
- **Calculate a z-score** z = 2
- In the table, a z-score of 2 represents a probability up to that point of .4772
- But we want the area after 2 standard deviations
- .5 - .4772 = .0228 one side of curve
- 2 x .0228 = .0456 both sides of curve
- Or 4.56% rounded to 5%

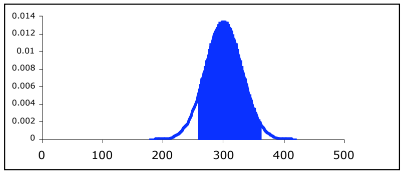


More than 3 standard deviations from the mean?

- X ~ N μ= 300 σ= 30
- More than 3 std deviations, z = 3.00
- In table when z = 3.00 we have a probability up to that point on one side of the curve of .4987
- .5 - .4987 = .0013 one side of curve
- 2 x .0013 = .0026 both sides of curve
- **.26% of the values are greater than 3 standard deviations from the mean**

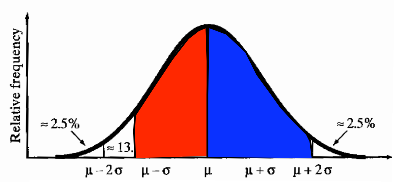
Problem: Probability that x is between 260 and 360

- X ~ N μ= 300 σ= 30
- Probability that x is between 260 and 360?
  - Draw it
  - Calculate z-scores
  - Look up in the table
  - Do any calculations



Probability that x is between 260 and 360?

- X =260 z = (260 – 300)/30 = **-1.33**
- X = 360 z = (360 – 300)/30 = **2.00**
- Since the table shows only one side, use absolute value
- The probability for z = 1.33 = **.4082**
- The probability for z = 2.00 = **.4772**
- The solution in this case is to add the two probabilities
  - **.4082 + .4772 = .8854**



What is the value at the 80th percentile?

- What am I looking for?
- I am looking for the X value that corresponds to the 80th percentile
- The 80th percentile reflects everything up to the mean (50th percentile)
- Plus .30 more
- For this problem I am looking for a probability of .30 inside the table, and reading out to the z-score
- Why?

$$z = \frac{(X - \mu)}{\sigma}$$

What is the value at the 80th percentile?

- Look inside the table for p = .30
- It is between z=.84 (p = .2995) and z=.85 (p= .3023)
- I could extrapolate, but I know it is a lot closer to z= .84
- z= .842 is a good approximation
- So now I can solve for the value of X that corresponds to a z-value

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621

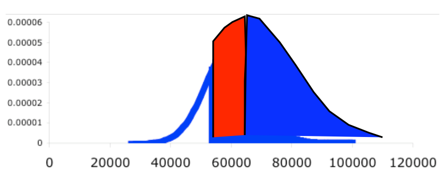
The 80th percentile for our normally distributed variable with μ=300 and σ= 30 is at 325.26

You solve it

- The tread life of a particular brand of tire is a random variable best described by a normal distribution with a mean of 60,000 miles and a standard deviation of 8,300 miles.
- **What proportion of the tires last longer than 53,775 miles?**
- STEPS:
  - Draw it out
  - Calculate z-score
  - Check the table
  - Do any final calculations

Solution: proportion of the tires last longer than 53,775 miles

- **Draw it out**
- **Calculate:** z = (53,775 – 60,000)/8,300 = **-.75**
- The probability associated with z= .75 is .2734 – this is the part on the left side up to 60,000 miles
- **Further Calculations:** But I also have to include the right side of the distribution, the part after 60,000 miles!
- P = **.2734** + **.5** =
- **= .7734**

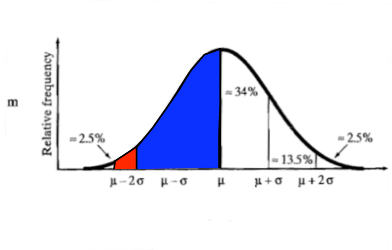


You solve it

- The tread life of a particular brand of tire is a random variable best described by a normal distribution with a mean of 60,000 miles and a standard deviation of 8,300 miles.
- **What proportion of tires last between 40,000 and 45,000 miles?**
- STEPS:
  - Draw it out
  - Calculate z-score
  - Check the table
  - Do any final calculations

Solution: proportion of the tires last between 40,000 and 45,000 miles

- **Draw it out**
- **Calculate:** z = (40,000 – 60,000)/8,300 = **-2.41**
- **Calculate:** z = (45,000 – 60,000)/8,300 = **-1.81**
- The probability associated with z = -2.41 and -1.81 are:
  - .4920
  - .4649
- **Further Calculations:** I want the part between then, so I subtract the two probabilities P = .4920 - .4649 = .0271
- **= .0271**

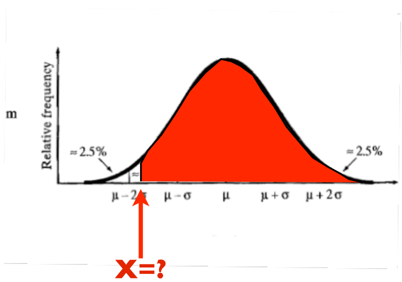


You solve it

- The tread life of a particular brand of tire is a random variable best described by a normal distribution with a mean of 60,000 miles and a standard deviation of 8,300 miles.
- **What warranty should the company use if they want 96% of the tires to outlast the warranty?**
- STEPS:
  - Draw it out
  - Calculate z-score
  - Check the table
  - Do any final calculations

Solution: What warranty should the company use if they want 96% of the tires to outlast the warranty?

- **Draw it out**
  - I want to find a value for X on the left-hand side which corresponds to a probability of .96 to the right
  - It is at the 4th percentile
- Since my table represents 1/2 of the curve, I want a probability of **.46**
- One probability is close, .4599, which corresponds to a **z = 1.75**
- And I want this to be **-1.75**
- **Calculate:** solve for the value of X at the 4th percentile
- **Answer: 45,475**



For Proportions, use the Normal Distribution as an approximation of the Binomial Distribution

- Proportions can be thought of as coming from a binomial distribution
  - Binomial: # success/total with a constant p of success
- Suppose we conducted a survey of 100 people and 56 answered yes to a question on whether they intended to vote
  - **Mean:** p_{yes} = #Yes/Total = 56/100 = **.56**
  - **Variance:** p*(1-p) = p*q = .56*.44 = **.2464**

For p = .56, the binomial looks like a normal when n is large

- For n=100, the binomial distribution looks like a normal distribution
- For n=50, this still holds true
- For n= 8, it looks less like a normal distribution
- And the more extreme p is, meaning the closer to zero or 1, the worse things get
  - p = .20, n= 10
  - p = .10, n= 30

To use the Normal Approximation to the Binomial Distribution (for proportions)

- It is ok to use the normal approximation whenever both
  - n*p > 5 and n*q > 5
- Example: n = 50 p = .2 and q = .8
  - 50*.2 = 10
  - 50*.8 = 40

YES!! ☺

- Example: n = 50 p = .05 and q = .95
  - 50*.05 = 2.5
  - 50*.95 = 47.5

NO!! ☹

  - **I would need n > 100 to make this work**

To use the Normal Approximation to the Binomial Distribution (for proportions)

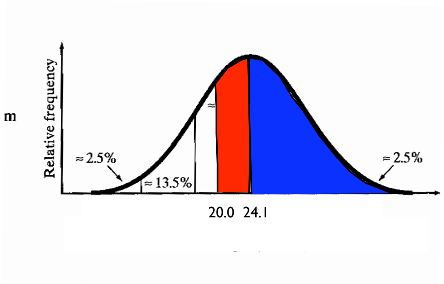
- For our purposes we will note that when n is reasonable large (n > 50), and p or q is not extremely small (p and q > .10), we can generally use this approach, since:
  - 50*.10 = 5
- Some books suggest a Continuity Correction of adding .5 to the value of X (the # of successes).
- This is because the binomial refers to a discrete random variable and the Normal Distribution is continuous

You Solve It

- The physical fitness of a patient is measured by the maximum oxygen uptake (recorded in milliliters per kilogram, ml/kg)
- The maximum oxygen uptake for cardiac patients, **CardOup**, who regularly participate in sports or exercise programs was found to be:
- **CardOup** ~N(24.1, 6.3)
- **What is the probability that a cardiac patient who regularly participates in sports has a maximum oxygen uptake of at least 20 ml/kg?**

Answer: probability that a cardiac patient who regularly participates in sports has a maximum oxygen uptake of at least 20 ml/kg?

- At least 20 means
  - 20 up to the mean of 24.1
  - Plus everything past 24.1
- z = (20 – 24.1)/6.3 = **-.65**
- Normal Table p = .2422
- **Answer = .2422 + .5 = .7422**

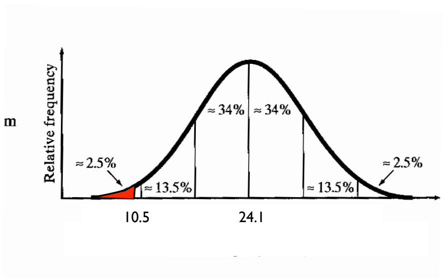


You Solve It

- The physical fitness of a patient is measured by the maximum oxygen uptake (recorded in milliliters per kilogram, ml/kg)
- The maximum oxygen uptake for cardiac patients, **CardOup**, who regularly participate in sports or exercise programs was found to be:
- **CardOup** ~N(24.1, 6.3)
- **What is the probability that a cardiac patient who regularly participates in sports has a maximum oxygen uptake of 10.5 ml/kg or lower?**

Answer: probability that a cardiac patient who regularly participates in sports has a maximum oxygen uptake of 10.5 ml/kg or lower?

- At least 10.5 or lower means the area in the left tail, after 10.5
- Calculate the z-score for 10.5
  - z = (10.5 - 24.1)/ 6.3
  - z = **-2.16**
- Normal Table p = .4846
- But we want the area to the left of 10.5
- **Answer = .5 - .4846 = .0154**
- **1.54%**



Rare Event?

- Consider a cardiac patient with a maximum oxygen uptake of 10.5 ml/kg.
- Is it likely that this patient participates regularly in sports or exercise programs?
- The way we think of this is, “what is the probability for this value out into the left tail?”
- If it is small, then this is a relatively rare event. So rare, that we might cast doubt on whether this patient participates in sports or exercise programs.
- **In our case, p = .0154 This is a rare event.**

Summary

- Be familiar with the normal distribution and table gymnastics
- When we shift to inference, the normal distribution and the related t-distribution, will be very important
- We will couple this with the rare event approach to begin to construct confidence intervals and conduct hypothesis tests