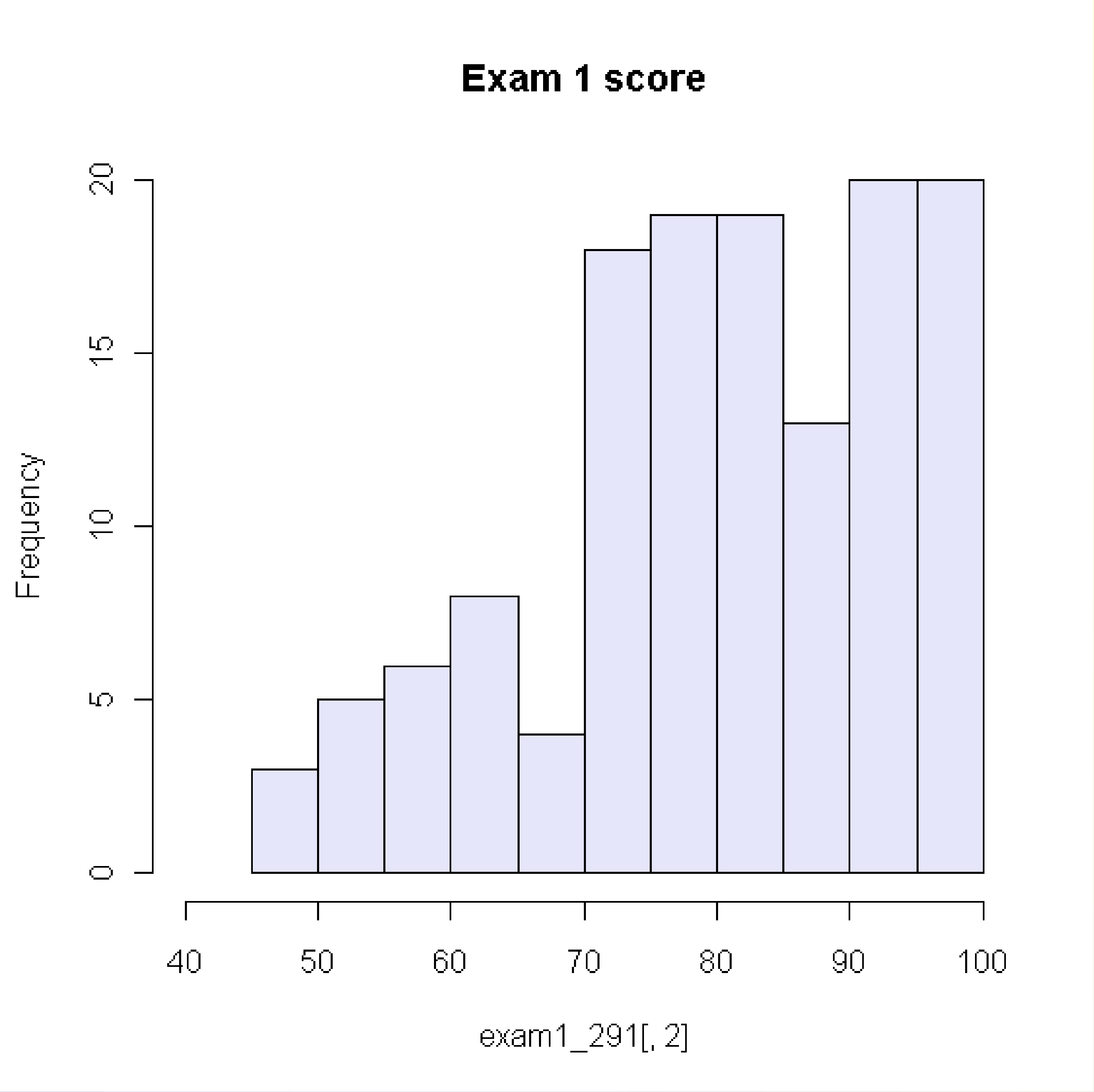
# STA 291
## Lecture 15

– **Normal Distributions (Bell curve)**


Exam 1 score
Frequency
0
5
10
15
20
40
50
60
70
80
90
100
exam1_291[, 2]

- Mean = 80.98
- Median = 82
- SD = 13.6

- Five number summary:

- 46 74 82 92 100

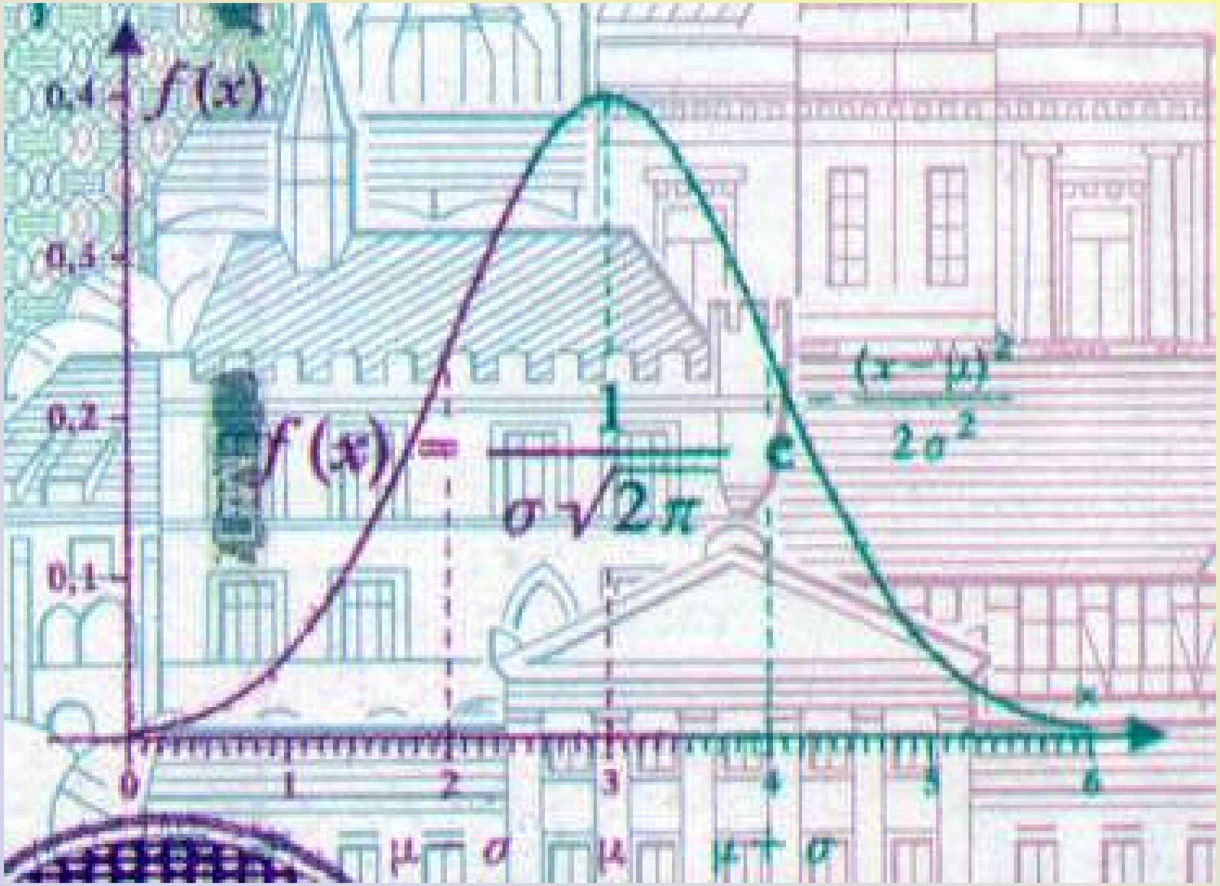
- There are many different shapes of continuous probability distributions…

- We focus on one type – the Normal distribution, also known as Gaussian distribution or bell curve.

# Carl F. Gauss


GL7067146A5
DEUTSCHE BUNDESBANK
10
ZEHN DEUTSCHE MARK
GL7067146A5

# Bell curve

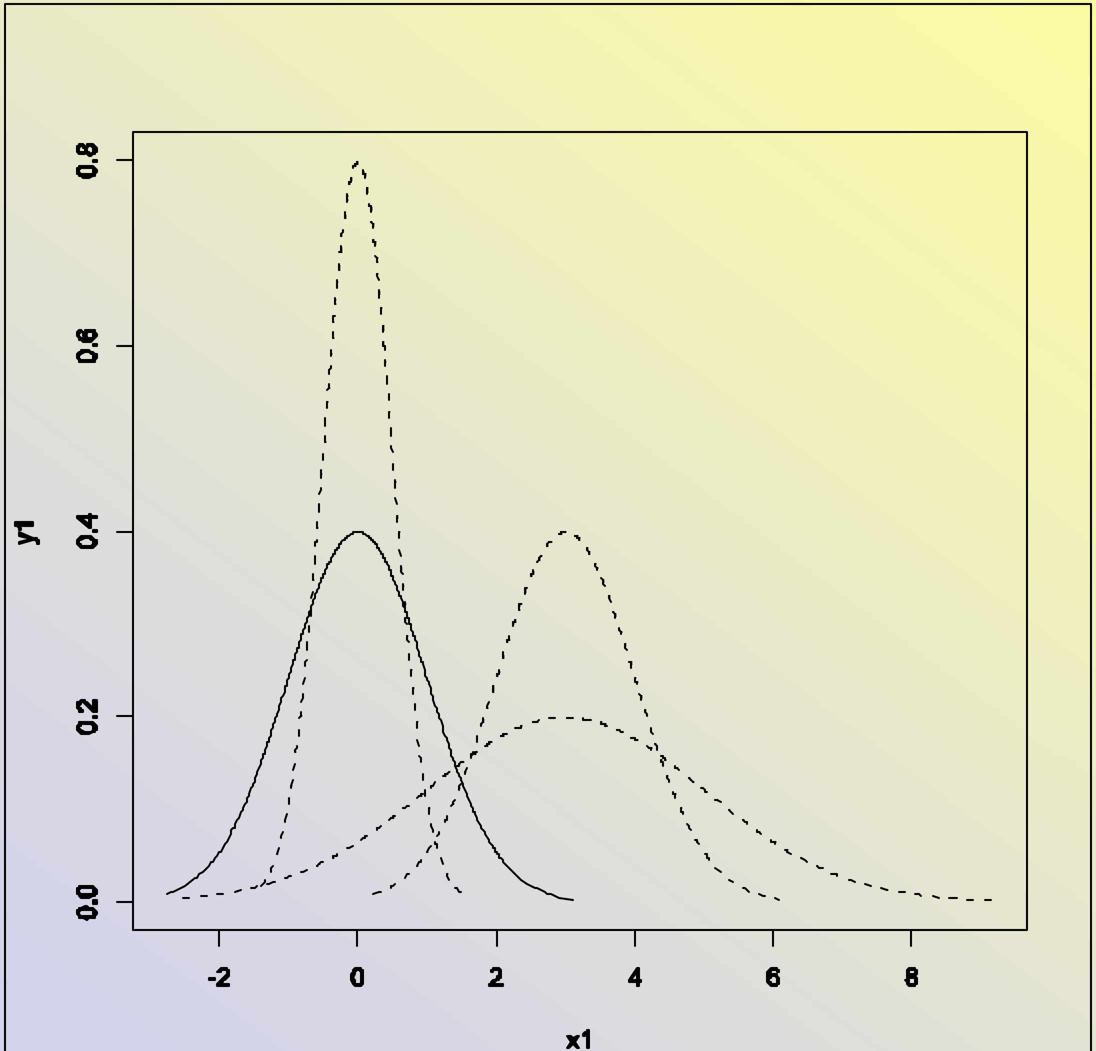


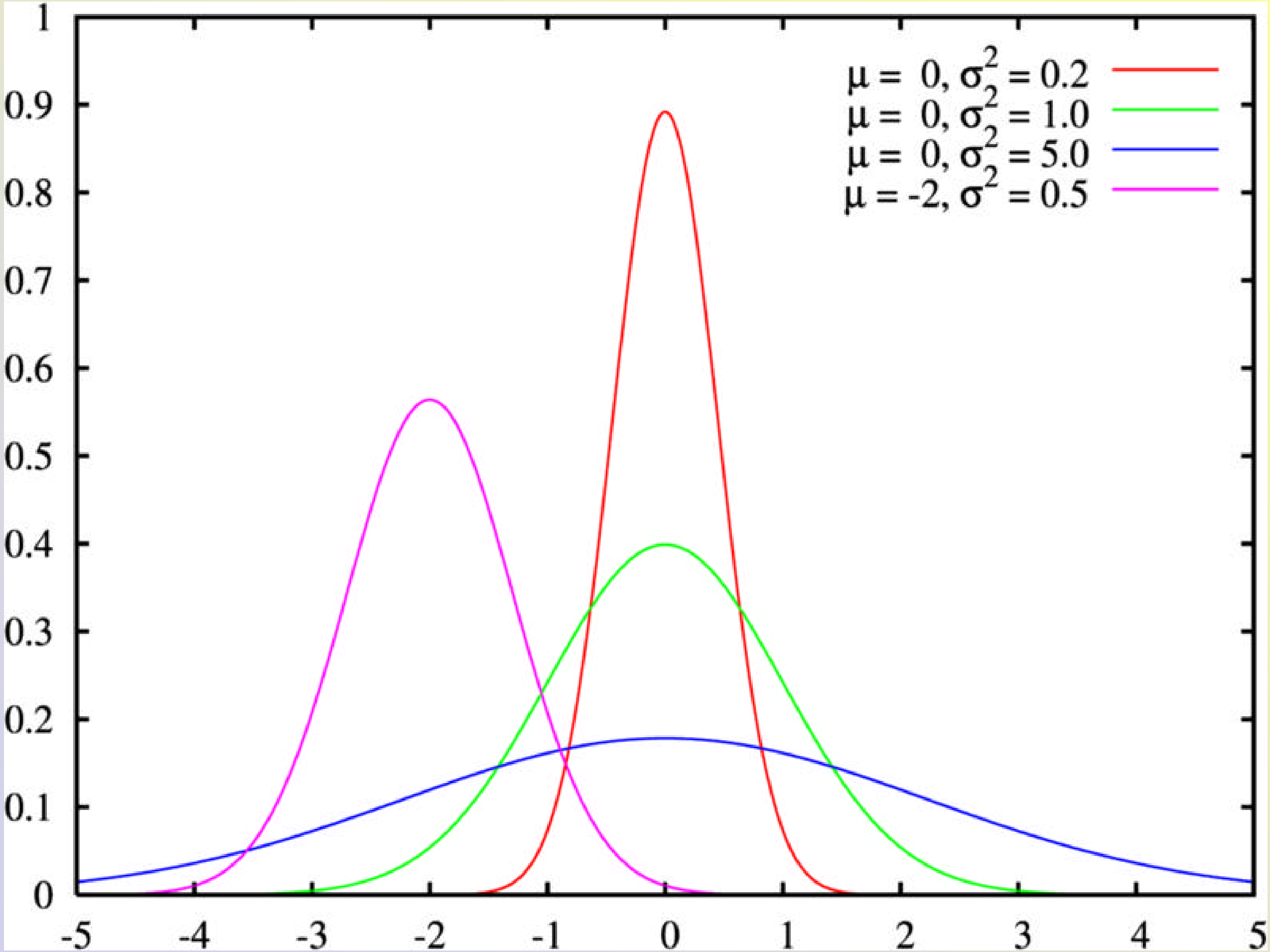
0,4
$f(x)$
0,3
0,2
0,1
$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
0
1
2
3
4
5
6
$\mu - \sigma$
$\mu$
$\mu + \sigma$

# Normal distributions/densities

- Again, this is a whole family of distributions, indexed by mean and SD. (location and scale)

# Different Normal Distributions





$\mu = 0, \sigma^2 = 0.2$
$\mu = 0, \sigma^2 = 1.0$
$\mu = 0, \sigma^2 = 5.0$
$\mu = -2, \sigma^2 = 0.5$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-5
-4
-3
-2
-1
0
1
2
3
4
5

# The Normal Probability Distribution

- Normal distribution is perfectly ***symmetric*** and ***bell-shaped***
- Characterized by **two parameters**:

***mean μ*** and ***standard deviation s***

- The ***68%-95%-99.7%*** rule applies to the normal distribution.
- That is, the probability concentrated within 1 standard deviation of the mean is always 68% within 2 SD is 95%; within 3 SD is 99.7% etc.

- It is very common.

- The sampling distribution of many common statistics are approximately Normally shaped, when the sample size n gets large.

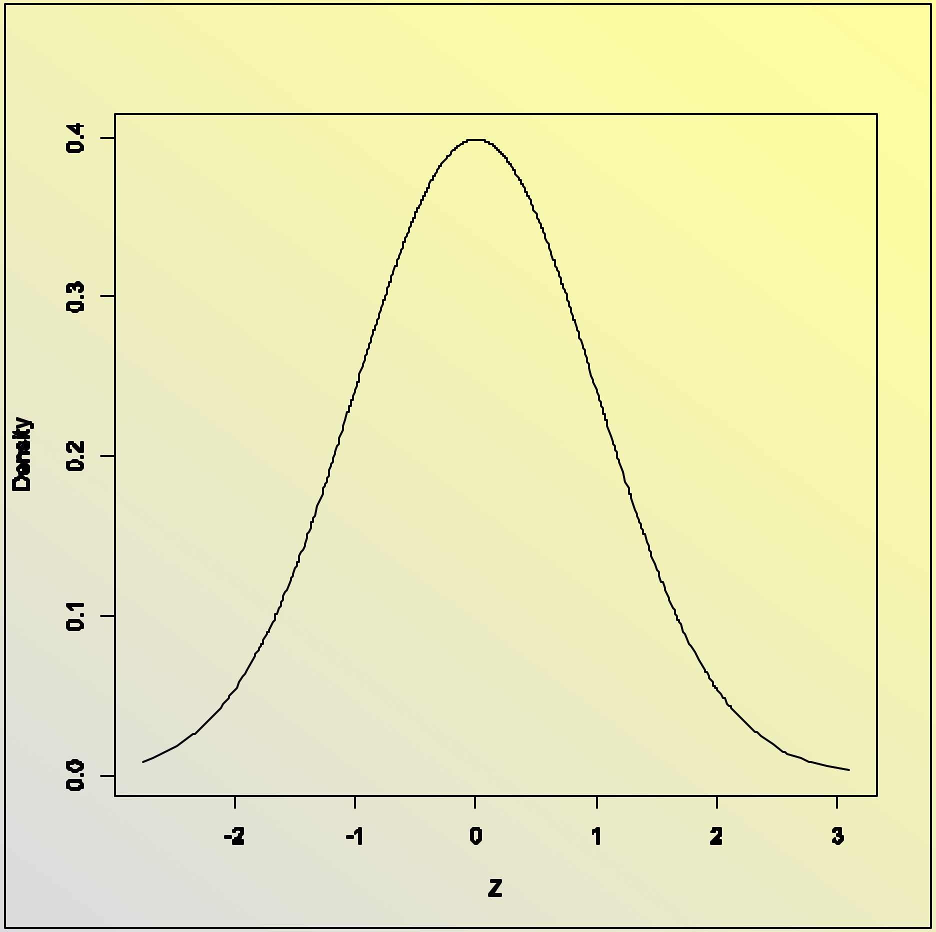
# Central Limit Theorem

- In particular:
- Sample proportion $\hat{p}$
- Sample mean $\bar{X}$
- The sampling distribution of both will be approximately Normal, for large n

# Standard Normal Distribution

- The standard normal distribution is the normal distribution with mean μ=0 and standard deviation

$$s = 1$$



# Non-standard normal distribution

- Either mean $m \neq 0$
- Or the SD $s \neq 1$
- Or both.
- In real life the normal distribution are often non-standard.

# Examples of normal random variables

- Public demand of gas/water/electricity in a city.
- Amount of Rain fall in a season.
- Weight/height of a randomly selected adult female.

# Examples of normal random variables – cont.

- Soup sold in a restaurant in a day.
- Stock index value tomorrow.

# Example of *non-normal* probability distributions

Income of a randomly selected family.
(skewed, only positive)

Price of a randomly selected house.
(skewed, only positive)

# Example of non-normal probability distributions

- Number of accidents in a week. (discrete)

- Waiting time for a traffic light. (has a discrete value at 0, and only with positive values, and no more than 3min, etc)

# Central Limit Theorem

- Even the incomes are not normally distributed, the ***average*** income of many randomly selected families ***is*** approximately normally distributed.

- Average does the magic of making things normal!          (transform to normal)

# Table 3 is for standard normal

- Convert non-standard to standard.
- Denote by X -- non-standard normal
- Denote by Z -- standard normal

# Standard Normal Distribution

- When values from an arbitrary normal distribution are converted to *z*-scores, then they have a standard normal distribution
- The conversion is done by subtracting the mean µ, and then dividing by the standard deviation s

$$z = \frac{x - m}{s}$$

# Example

- Find the probability that a randomly selected female adult height is between the interval 161cm and 170cm. Recall

$$m = 165, s = 8$$

$$\frac{161 - 165}{8} = -0.5;$$

$$\frac{170 - 165}{8} = 0.625$$

# Example –cont.

- Therefore the probability is the same as a standard normal random variable Z between the interval -0.5 and 0.625

$$P(161 < X < 170) = P(-0.5 < Z < 0.625)$$

# Use table or use Applet?

- Applet 1: http://bcs.whfreeman.com/scc/content/cat_040/spt/normalcurve/normalcurve.html

- Applet 2:

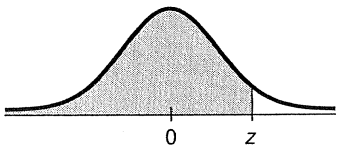
http://psych-www.colorado.edu/~mcclella/java/normal/handleNormal.html

# Online Tool

- Normal Density Curve
- Use it to verify graphically the empirical rule, find probabilities, find percentiles and *z*-values for one- and two-tailed probabilities

**Table Z (cont.)**

*Areas under the standard Normal curve*




| | Second decimal place in z | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |

# *z*-Scores

- The *z*-score for a value x of a random variable is the number of standard deviations that x is above μ
- If x is below μ, then the *z*-score is negative
- The *z*-score is used to compare values from different normal distributions

# Calculating *z*-Scores

- You need to know x, µ, and $s$
  to calculate *z*

$$z = \frac{x - m}{s}$$

- Applet does the conversion automatically. (recommended)

The table 3 gives probability

$$P(0 < Z < z) = ?$$

# Tail Probabilities

- SAT Scores: Mean=500,

  SD =100
- The SAT score 700 has a *z*-score of *z*=2
- The probability that a score is ***beyond*** 700 is the tail probability of Z ***beyond*** 2

# $z$-Scores

- The $z$-score can be used to compare values from different normal distributions
- SAT: μ=500, s=100
- ACT: μ=18, s=6
- Which is better, 650 in the SAT or 26 in the ACT?

$$z_{SAT} = \frac{x - m}{s} = \frac{650 - 500}{100} = 1.5$$

$$z_{ACT} = \frac{x - m}{s} = \frac{26 - 18}{6} = 1.333$$

- Corresponding tail probabilities?
- How many percent of total test scores have better SAT or ACT scores?

# Typical Questions

1. Probability (right-hand, left-hand, two-sided, middle)
2. z-score
3. Observation (raw score)

- To find probability, use applet or Table 3.
- In transforming between 2 and 3, you need mean and standard deviation

# Finding *z*-Values for Percentiles

- For a normal distribution, how many standard deviations from the mean is the 90th percentile?
- What is the value of $z$ such that 0.90 probability is less than $\mu + z\, s$ ?
- If 0.9 probability is less than $\mu + z\, s$, then there is 0.4 probability between 0 and $\mu + z\, s$ (because there is 0.5 probability less than 0)
- $z$=1.28
- The 90th percentile of a normal distribution is 1.28 standard deviations above the mean

# Quartiles of Normal Distributions

- Median: $z=0$

  (0 standard deviations above the mean)
- Upper Quartile: $z = 0.67$

  (0.67 standard deviations above the mean)
- Lower Quartile: $z= -\ 0.67$

  (0.67 standard deviations below the mean)

- In fact for any normal probability distributions, the 90th percentile is always

1.28 SD above the mean

the 95th percentile is ____ SD above mean

# Finding $z$-Values for Two-Tail Probabilities

- What is the $z$-value such that the probability is 0.1 that a normally distributed random variable falls more than $z$ standard deviations ***above or below*** the mean
- Symmetry: we need to find the $z$-value such that the right-tail probability is 0.05 (more than $z$ standard deviations ***above*** the mean)
- $z$=1.65
- 10% probability for a normally distributed random variable is outside 1.65 standard deviations from the mean, and 90% is within 1.65 standard deviations from the mean

# homework online

# Attendance Survey Question 16

- On a 4”x6” index card
  - Please write down your name and section number
  - Today’s Question:

___?___ is also been called bell curve.