Variables

Numerical

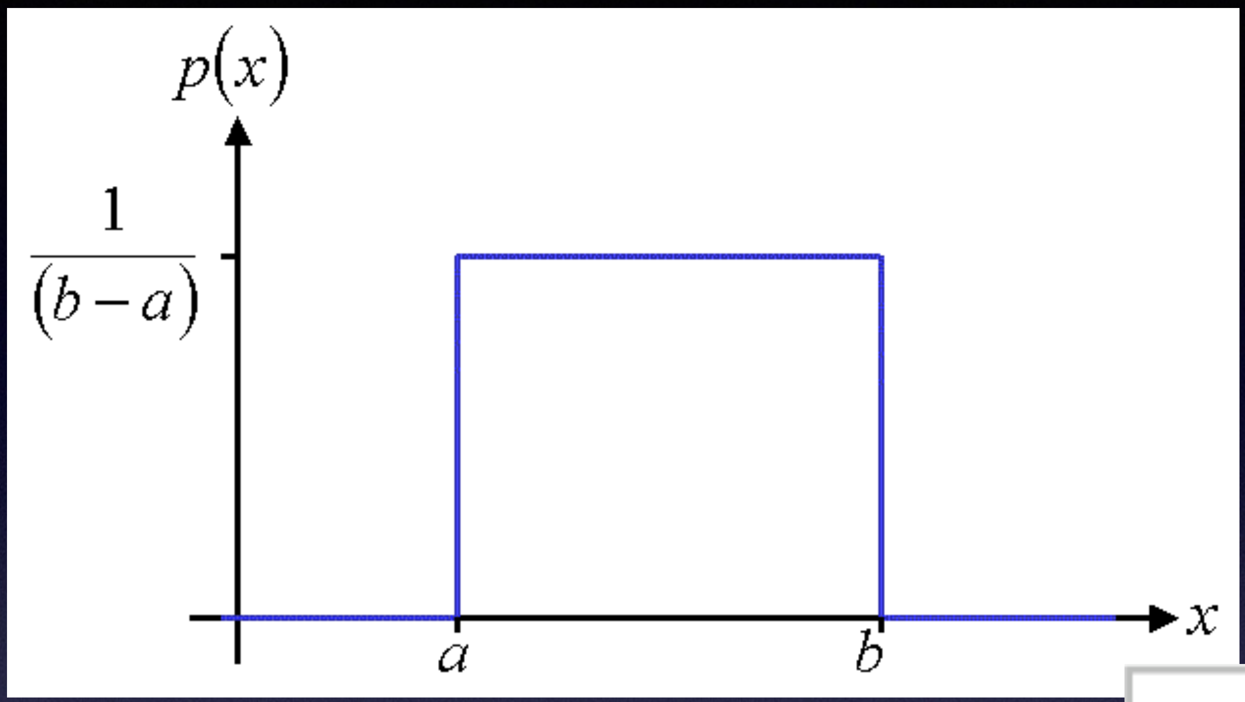
Categorical

Numeric
•Discrete
•Continuous

Nonnumeric

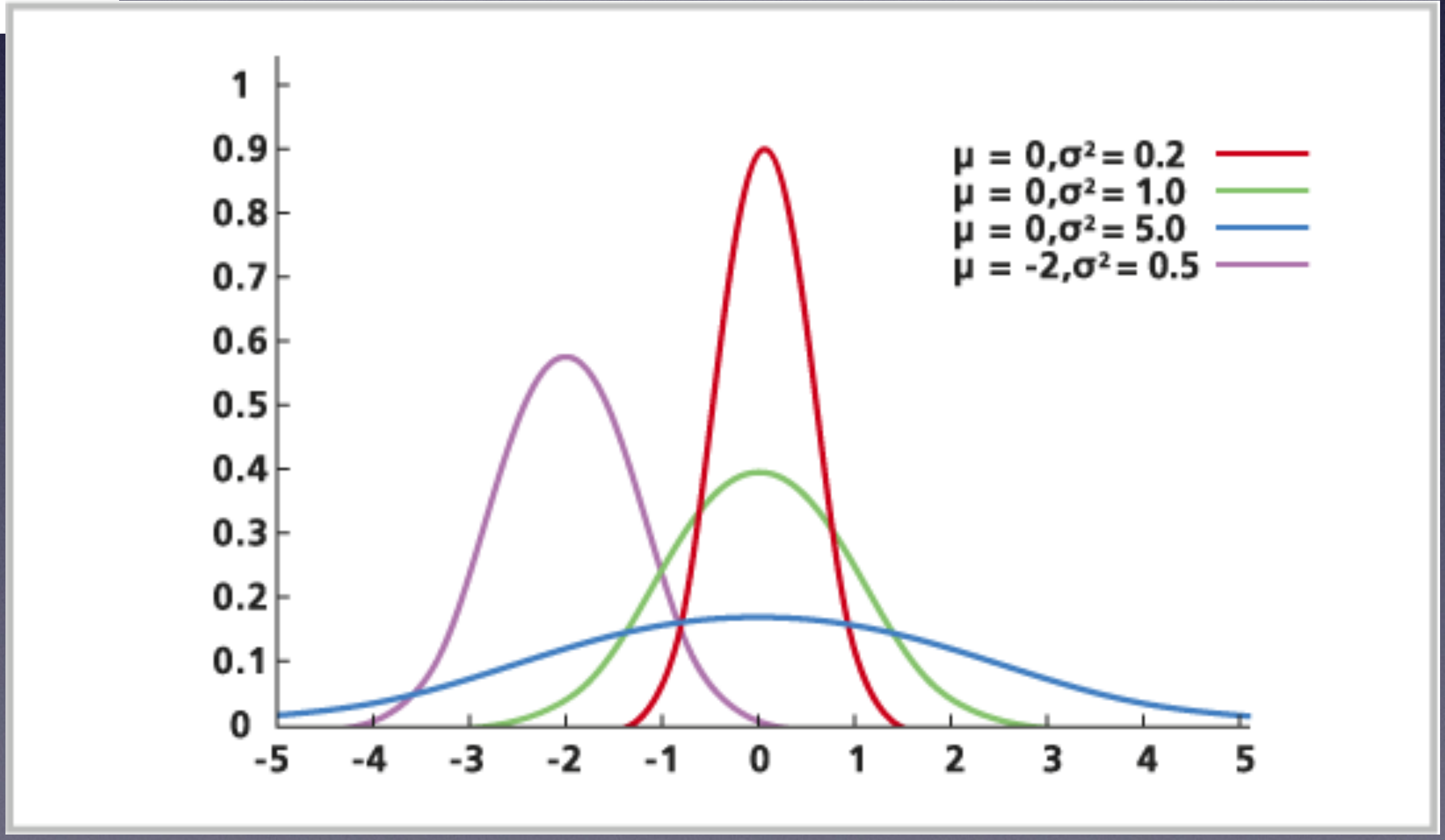
We will first focus on Continuous Random Variables

Most Common Continuous RV Distributions -



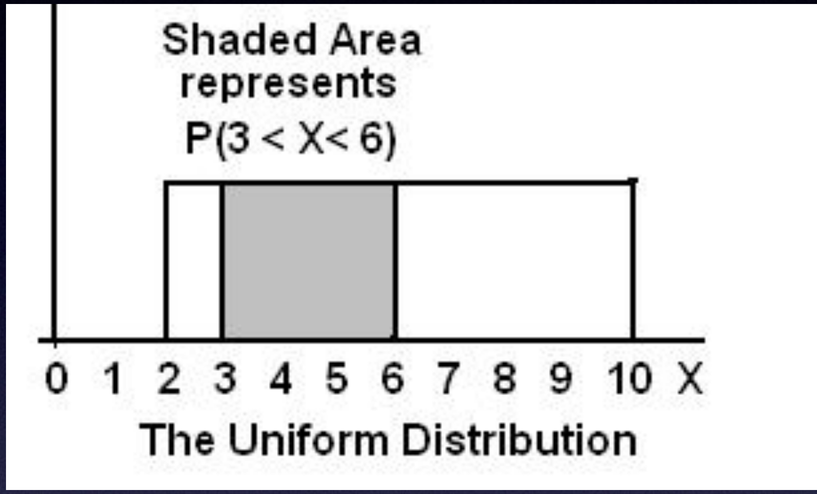
Uniform Distribution

Normal Distribution

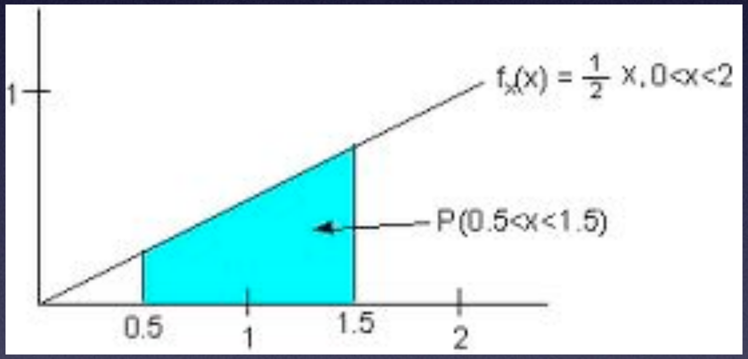


Probability for Continuous RVs

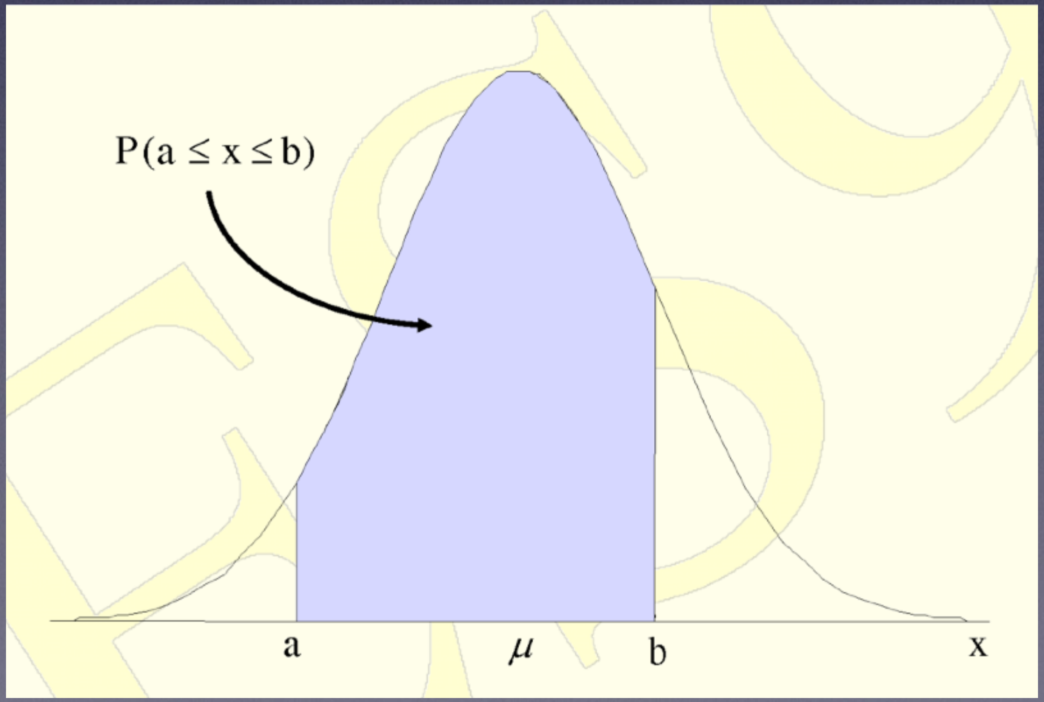
= AREA UNDER A CURVE



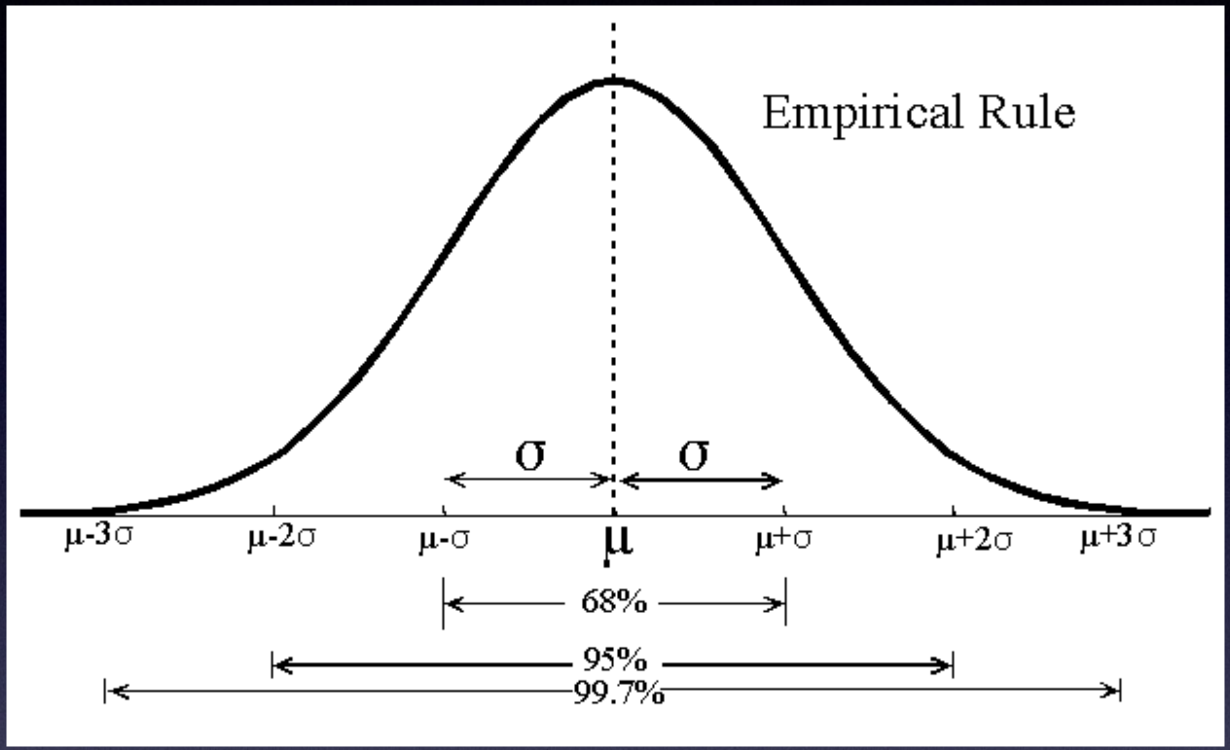
Uniform Distribution



Normal Distribution



Empirical Rule - the 68/95/99.7 Rule



Approximately 68% of the observations are within 1 standard deviation of the mean. (z-score = ±1)

Approximately 95% of the observations are within 2 standard deviation of the mean. (z-score = ±2)

Approximately 99.7% of the observations are within 3 standard deviation of the mean. (z-score = ±3)

Deals with the ***middle*** *___ % of the data*

Percentiles - value such that ___ % of the observations in the data set fall ***below*** that value

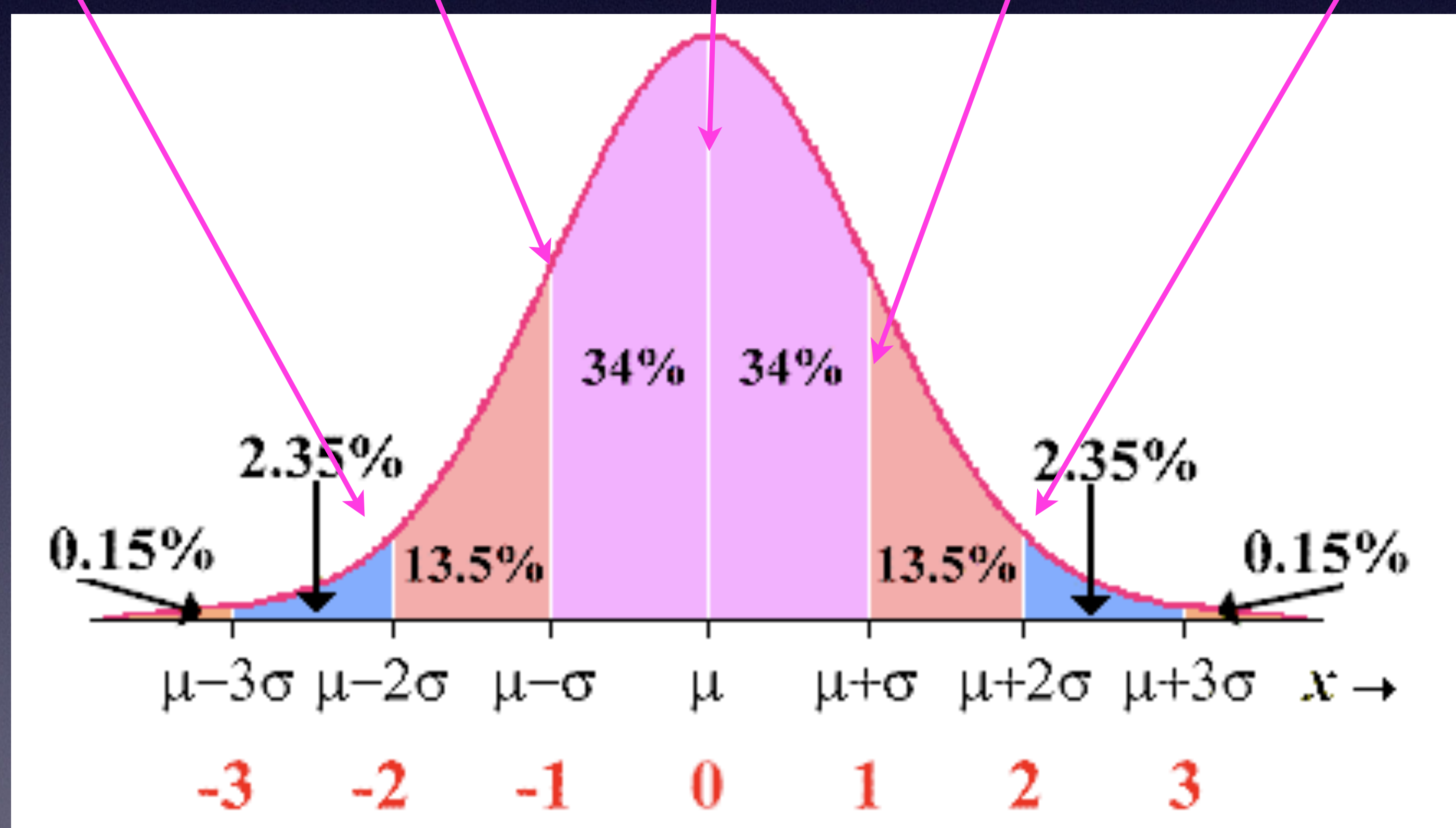
2.5th %ile

50th %ile

97.5th %ile

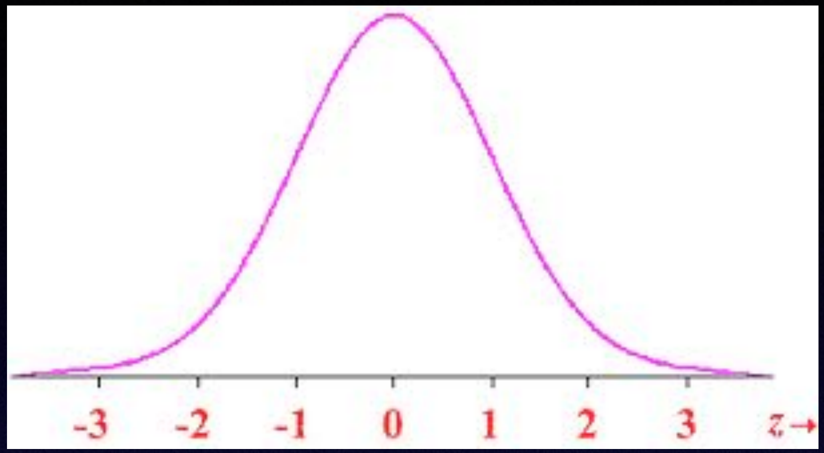
16th %ile

84th %ile



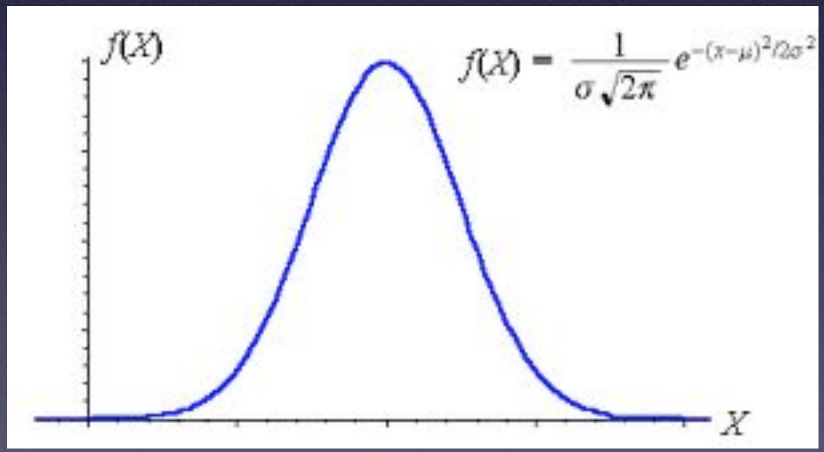
Standard Normal Distribution

A normal distribution in which the mean is 0, the standard deviation is 1, and $x = z$ (the x value is equal to the z score).



$$z \sim N(0,1)$$

$$P(z > \#) = normalcdf(\#,\ 1E99,\ 0,\ 1)$$



Normal Distribution

$$x \sim N(\mu,\sigma)$$

$$P(x < \#) = normalcdf(-1E99,\ \#,\ \mu,\ \sigma)$$

This is graphing calculator stuff. You will soon see…

z scores

$$= \frac{\text{value} - \text{mean}}{\text{standard deviation}}$$

- Provide a common scale to compare data
- Conveys how many standard deviations above/below the mean a data value is
- Positive z scores lie above the mean
- Negative z scores lie below the mean

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

Now bear with me while we do a little algebra

$$z\sigma = x - \mu$$

$$x = \mu + z\sigma$$

Why did we do this?
Let's find out.

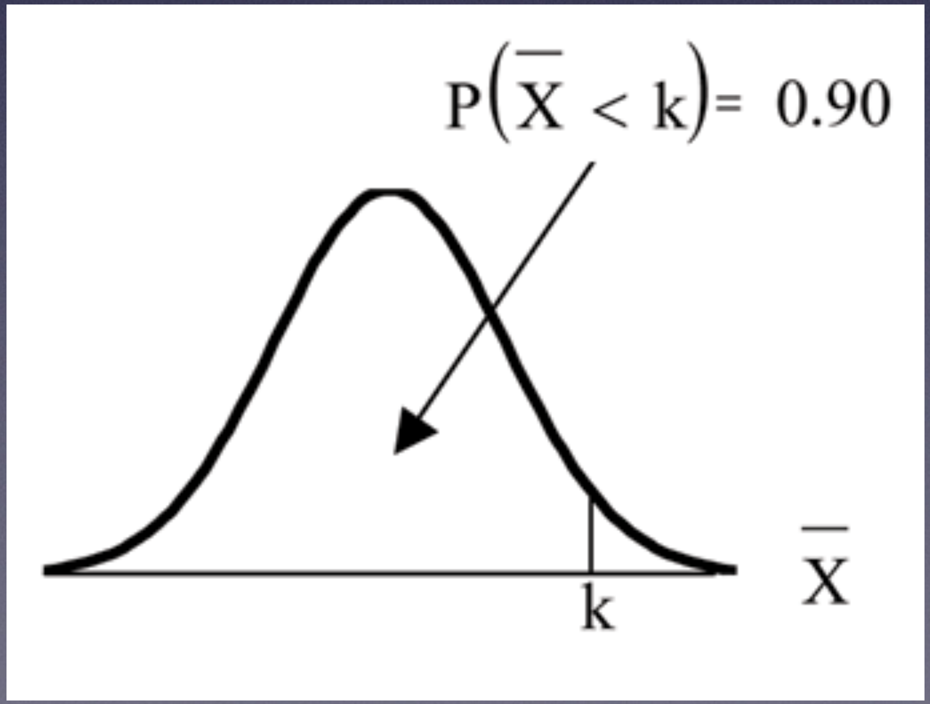
The Backwards Problems - i.e. invNorm

- Calculator $x = invNorm(\%\text{ile},\ \mu,\ \sigma)$

- Formula - not on formula sheet

$$x = \mu + z\sigma$$

You can leave these two blank if you have a standard normal distribution



$$k = invNorm(0.90,\ \mu,\ \sigma)$$

The Backwards Problems - i.e. invNorm

- Calculator $x = invNorm(\%\text{ile}, \mu, \sigma)$

- Formula - not on formula sheet

$$x = \mu + z\sigma$$

A very, very, very, very important thing about invNorm…

You can use it to find a *z* score when you only know the percentile

$$z = invNorm(0.90) = 1.28155$$

This means that the *z* score of the 90th percentile in ***any*** normal distribution is 1.28155

Try these out on your calculator because it’s a skill you will need. If you need to see further demos, watch the next screencast in this unit.