# MATH 124 Spring 2005

# Lecture: 18
# Date: Apr 21, 2005

**Skills you should acquire from this lecture:**

- Dealing with sampling distribution of the sample mean, examples and worked through problems
- Finding the mean, variance and standard deviation of linear combinations of random variables.

**Related readings in the textbook:**

- Section 5.2
- Problems 5.39, 5.61 (Discussed at TTh 8:10-9:25)
- Problems 5.38, 5.40 (Discussed at TTh 9:35-10:50)

# Example

A machine produces parts for a car. This part has mean weight 150g and, because the machine has variability, standard deviation 10g. Each hour a SRS of 100 of the parts is taken to make sure the machine is still producing quality parts. Suppose that an entire hours output will have to be discarded if the mean weight of the sample is below 147.5g. What is the probability of this happening?

$$\mu_{\bar{x}} = 150g \quad \sigma_{\bar{x}} = \frac{10}{\sqrt{100}} = \frac{10}{10} = 1g$$

because $n = 100$ is large the CLT tells us that $\bar{x}$ is approximately Normally distributed with mean 150 and stddev 1g.

approximately equal due to CLT

$$P(X < 147.5) \approx P\left(\frac{X - 150}{1} < \frac{147.5 - 150}{1}\right)$$

$= P(Z < -2.5)$


$= .0062$ (From table)

Example

An experiment to compare the nutrional value of two feeding regimins for chickens is carried out on 40 chicks. At random 20 are assigned to diet 1 and 20 to diet 2. At the end of the experiment the weight gains are measured. Call $\bar{x}$ the mean weight gain for diet 1 and $\bar{y}$ the mean weight gain for diet 2. At the end of experiment inference about which diet is better will be based on the difference $\bar{y}-\bar{x}$.

Suppose $\mu_x = 360$ and $\sigma_x = 55$ and

$\mu_y = 385$ $\sigma_y = 50$.

What is the mean of $\bar{y}-\bar{x}$?

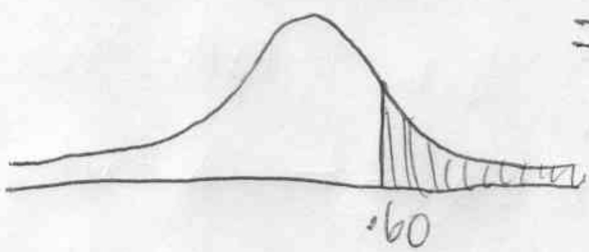
$\mu_{\bar{y}-\bar{x}} = \mu_{\bar{x}} - \mu_{\bar{y}} = 385 - 360 = 25g$

$\sigma_{\bar{y}-\bar{x}} = \sqrt{\sigma^2_{\bar{y}} + \sigma^2_{\bar{x}}} = \sqrt{\frac{55^2}{20} + \frac{50^2}{20}} = 16.6208$

Assuming that $\bar{x}$ and $\bar{y}$ are normally distributed then $\bar{y} - \bar{x}$ is normally distributed with mean 25g and std dev 16.6208

what is the probability that diet 2 has average weight gain of more than 35g above that of diet 1?

$$P(\bar{y} - \bar{x} \geq 35) = P\left(\frac{(\bar{y} - \bar{x}) - 25}{16.6208} \geq \frac{35 - 25}{16.6208}\right)$$



$$= P(Z \geq .60)$$

$$= 1 - P(Z < .60)$$

$$= 1 - .7257$$

$$= .2743$$

# Problem 5.39

Central limit theorem tells us that as sample size $n$ increases the sampling distribution of $\bar{x}$ approaches normal with mean $\mu$ and standard deviation $\frac{\sigma}{\sqrt{n}}$

since $\mu = 1.6$, $\sigma = 1.2$ $n = 200$

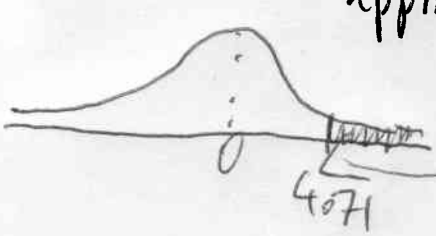
$\bar{x}$ has approximately normal distribution with mean 1.6, standard deviation $\frac{1.2}{\sqrt{200}} = 0.0849$

Probability statement

$$P(\bar{x} > 2) \approx P\Big(\underbrace{\frac{\bar{x} - 1.6}{.0849} > \frac{2 - 1.6}{.0849}}_{\text{standardization}}\Big)$$

normal approximation

$$= P(Z > 4.71)$$




note this is outside the limits of table.

$$= 1 - P(Z < 4.71)$$

$$\ll 1 - P(Z < 3.49)$$

largest value in table

$$\ll .0002$$

# Problem 5.61

(a) Not binomial because $n$ changes between vehicles (ie not every car has same size and same potential number of seats)

(b) Can't be normal because can't have half a person (nor negative people)

(c) Central limit theorem tells us that as $n$ increases the distribution of $\bar{x}$ approaches normal with mean 1.5 and standard deviation $\frac{.75}{\sqrt{700}} = .0283$

(d) Need mean and stddev of $700\bar{x}$

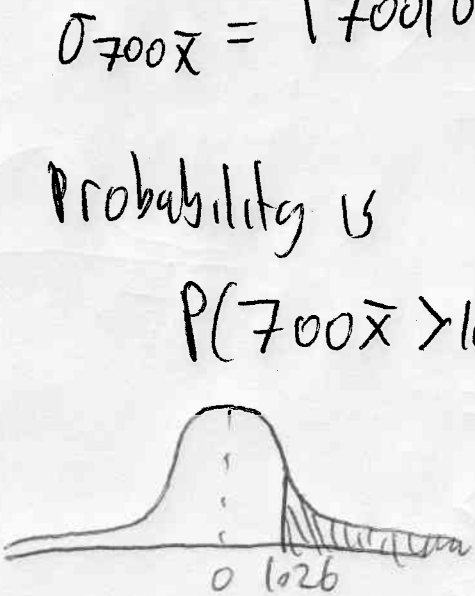
$$\mu_{700\bar{x}} = 700\mu_{\bar{x}} = 700(1.5) = 1050$$

$$\sigma_{700\bar{x}} = |700|\sigma_{\bar{x}} = 700\frac{(.75)}{\sqrt{700}} = 19.8431 \text{ (4dp)}$$

Probability is

approximation due to CLT

$$P(700\bar{x} > 1075) \approx P\left(\frac{700\bar{x} - 1050}{19.8431} > \frac{1075-1050}{19.8431}\right)$$




$$= P(Z > 1.26)$$

$$= 1 - P(Z < 1.26)$$

$$= 1 - .8962 = .1038$$

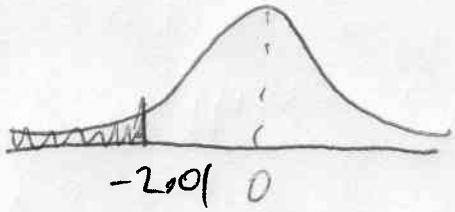
## Problem 5.38

(a) We know from the problem that X is exactly normally distributed with mean 55000 and standard deviation 4500. therefore the distribution of $\bar{X}$ is exactly normal with mean 55000 and standard deviation $\frac{4500}{\sqrt{8}} = 1590.99$

(b) $$P(\bar{x} \leq 51800) = \overbrace{P\left(\frac{\bar{x} - 55000}{1590.99} \leq \frac{51800 - 55000}{1590.99}\right)}^{\text{standardization step}}$$

$$= P(Z \leq -2.01)$$

$$= .0222$$



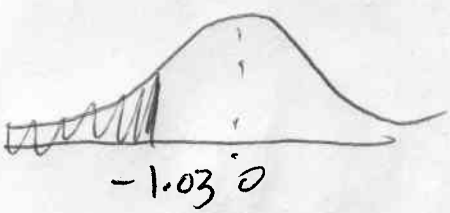

## Problem 5.40

(a) According to the CLT as n increases the distribution of $\bar{x}$ approaches normal with mean 2.2 and standard deviation $\frac{1.4}{\sqrt{52}} = .1941$ (4dp)

(b) $$P(\bar{x} < 2) \overset{\text{CLT approximation}}{\approx} P\left(\frac{\bar{X} - 2.2}{.1941} < \frac{2 - 2.2}{.1941}\right)$$

$$= P(Z < -1.03)$$

$$= .1515$$

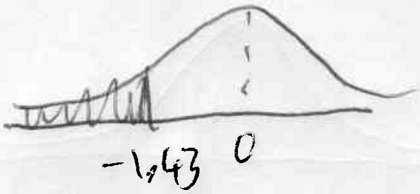



(c) Need mean and sd of $52\bar{x}$

$\mu_{52\bar{x}} = 52(2.2) = 114.4$ $\quad \sigma_{52\bar{x}} = (52)\frac{1.4}{\sqrt{52}} = 10.0955$

$$P(52\bar{x} < 100) \approx P\left(\frac{52\bar{x} - 114.4}{10.0955} < \frac{100 - 114.4}{10.0955}\right)$$

$$= P(Z < -1.43)$$




$$= .0764$$

# Rules for linear combinations of random variables

Suppose $X, Y, Z, \ldots$ are random variables and $a, b, c, \ldots$ are constants

## Means

$$\mu_{X+Y} = \mu_X + \mu_Y \qquad \mu_{X-Y} = \mu_X - \mu_Y$$

$$\mu_{aX+bY} = a\mu_X + b\mu_Y \qquad \mu_{aX-bY} = a\mu_X - b\mu_Y$$

$$\mu_{aX+b} = a\mu_X + b$$

$$\mu_{X+Y+Z} = \mu_X + \mu_Y + \mu_Z$$

## Variances

We consider only cases where $X, Y, Z$ etc are independent of each other (Your textbook explains non independent cases)

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y \qquad \sigma^2_{X+Y+Z} = \sigma^2_X + \sigma^2_Y + \sigma^2_Z$$

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y \qquad \sigma^2_{aX-bY} = a^2\sigma^2_X + b^2\sigma^2_Y$$

$$\sigma^2_{aX+bY} = a^2\sigma^2_X + b^2\sigma^2_Y$$

$$\sigma^2_{aX+b} = a^2\sigma^2_X$$

Note that to get rules for standard deviations just square root both sides

eg $\sigma_{X-Y} = \sqrt{\sigma^2_X + \sigma^2_Y}$