8. The Central Limit Theorem

The Penny Experiment (photos on tablet)

1) Graph X = age of penny
   Population is not mound-shaped

2) Take samples of size $n = 5$
   Graph $\bar{x}$ = mean age of 5 pennies
   $\bar{x}$ is closer to mound-shaped

3) Take samples of size $n = 25$
   Graph $\bar{x}$ = mean age of 25 pennies
   $\bar{x}$ is almost mound-shaped

Notice: The $\bar{x}$-values are more tightly clustered than the original population

The SD of the $\bar{x}$-values is called the <u>standard error</u> or SE

SE is smaller than the original σ

Central Limit Theorem

When sample size $n \geq 30$, $\bar{x}$-values are approx. normal with mean $= \mu$ and $SE = \frac{\sigma}{\sqrt{n}}$

Translation: If $n \geq 30$, use $z = \frac{\bar{x} - \mu}{(\sigma/\sqrt{n})}$

Works regardless of shape of the original population!

Ex: A large class has a test average of 72 with a SD of 8. Take a random sample of n tests. Find P(n tests average to > 75) if:

a) $n = 30$

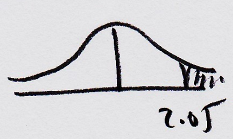
$P(\bar{x} > 75)$

$n \geq 30$ ✓

$z = \frac{\bar{x} - \mu}{(\sigma/\sqrt{n})} = \frac{75 - 72}{(8/\sqrt{30})} \approx 2.05$

$= P(z > 2.05)$

$=$



$= 0.5 - 0.4798 = 0.0202$

b) $n = 80$

$P(\bar{x} > 75)$

$n \geq 30$ ✓

$z = \dfrac{\bar{x} - \mu}{(\sigma/\sqrt{n})}$

$= \dfrac{75 - 72}{(8/\sqrt{80})}$

$\simeq 3.35$

$= P(z > 3.35)$

$=$

$= 0.5 - 0.4996$

$= 0.0004$

As sample size n gets bigger, $\bar{x}$-values stay closer to μ.

Ex: Checked baggage has a mean mass of 21 kg with a SD of 4 kg. 40 bags are randomly selected. Probability that average mass is:

a) between 20 and 23 kg?

$P(20 \leq \bar{x} \leq 23)$

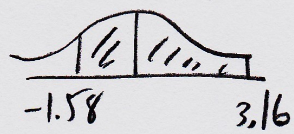
$n \geq 30$ ✓

$z_1 = \dfrac{20 - 21}{(4/\sqrt{40})} \simeq -1.58$

$z_2 = \dfrac{23 - 21}{(4/\sqrt{40})} \simeq 3.16$

$= P(-1.58 \leq z \leq 3.16)$

=



$= 0.4429 + 0.4992$

$= 0.9421$

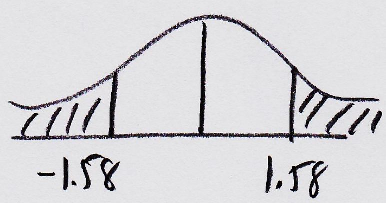
b) less than 20 kg or more than 22 kg?

$P(\bar{x} < 20 \text{ or } \bar{x} > 22)$

$$\begin{cases} n \geq 30 \checkmark \\ z_1 \approx -1.58 \\ z_2 = \dfrac{22-21}{(4/\sqrt{40})} \approx 1.58 \end{cases}$$

$= P(z < -1.58 \text{ or } z > 1.58)$

=



$= 1 - 2(0.4429)$

$= 0.1142$

<u>Ex</u>: Checked baggage mass

$\mu = 21\,kg \quad \sigma = 4\,kg$

P(total mass of 50 bags > 1130 kg)

total mass = 1130

average mass = $\frac{1130}{50}$

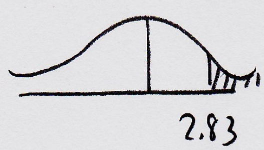
$\bar{x} = 22.6$

Want $P(\bar{x} > 22.6)$

$$\left\{\begin{array}{l} n \geq 30 \\ z = \dfrac{\bar{x} - \mu}{(\sigma / \sqrt{n})} \\ \quad = \dfrac{22.6 - 21}{(4/\sqrt{50})} \\ \quad \approx 2.83 \end{array}\right.$$

$= P(z > 2.83)$

$=$

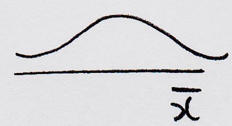


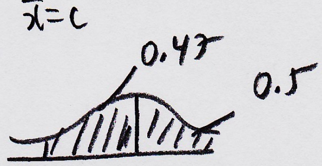
$= 0.5 - 0.4977$

$= \quad 0.0023$

<u>Ex</u>: Ball bearings have an average radius of 9.9mm with a SD of 1.4mm. Take a random sample of 60 ball bearings. Find c so that $P(\bar{x} > c) = 0.97$

$n \geq 30$ ✓





Reverse look-up area = 0.47
Use negative z

$z = -1.88$

$$z = \frac{\bar{x} - \mu}{(\sigma/\sqrt{n})} \qquad \text{Sub } \bar{x} = c,\ z = -1.88:$$

$$-1.88 = \frac{c - 9.9}{(1.4/\sqrt{60})}$$

$$-1.88\left(\frac{1.4}{\sqrt{60}}\right) = c - 9.9$$

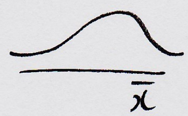
$$c = 9.9 - 1.88\left(\frac{1.4}{\sqrt{60}}\right) \approx 9.6 \text{ mm}$$

Ex: Ball bearing radius

$\mu = 9.9$ mm $\quad \sigma = 1.4$ mm

How large a sample would be required to ensure that $P(\bar{x} \geq 9.6) \geq 0.99$?

If $n \geq 30$:



shaded area = 0.99

$\bar{x} = 9.6$

$$\boxed{\begin{array}{l}\text{Reverse look-up area} = 0.49 \\ \quad \text{Use negative } z \\ \quad z = -2.33\end{array}}$$

$\bar{x} = 9.6$

$z = -2.33$

Want area ≥ 0.99

Need $z \leq -2.33$

$$\frac{\bar{x} - \mu}{(\sigma/\sqrt{n})} \leq -2.33$$

$$\frac{9.6-9.9}{(1.4/\sqrt{n})} \le -2.33$$

$$-0.3 \le -2.33\left(\frac{1.4}{\sqrt{n}}\right)$$

$$-0.3\sqrt{n} \le -2.33(1.4)$$

$$\sqrt{n} \ge \frac{-2.33(1.4)}{-0.3}$$ $\boxed{\text{Divided by negative \#}}$

$$n \ge \left[\frac{2.33(1.4)}{0.3}\right]^2$$

$$n > 118$$

RECAP:

1) If X is normal:

Use $z = \dfrac{x-\mu}{\sigma}$ where x = single measurement

2) If $n \ge 30$:

Use $z = \dfrac{\bar{x}-\mu}{(\sigma/\sqrt{n})}$ where $\bar{x}$ = sample mean = average of n measurements