Sample means are normally distributed:

If a variable itself is normally distributed then the distribution of sample means, $\overline{Y}$, is also normally distributed

Sampling distribution for $\overline{Y}$:

The range of different values for $\overline{Y}$ that could have been obtained by sampling, and their associated probabilities, constitute the sampling distribution for $\overline{Y}$.

Sample means are normally distributed:

If a variable itself is normally distributed then the distribution of sample means, $\bar{Y}$, is also normally distributed

- The mean of the sample means is μ
- The standard deviation of the sampling distribution for $\bar{Y}$ is called the *Standard error:*

$$\sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}} \xrightarrow{\text{approx}} SE_{\bar{Y}} = \frac{s}{\sqrt{n}}$$

Standard error **is the spread/variation statistic for a "collection" of means (or proportions)**.

It can be thought of as the theoretical variation present when you repeat a study many times.

This is in contrast to standard deviation which is used to describe the natural variation of something that you can measure

Standard Normal applied to sampling means:

• Sample means distributed normally with mean equal to μ and standard error then:

$$Z = \frac{\bar{Y} - \mu}{\sigma_{\bar{Y}}}$$

Practice Problem #6: Singleton babies born have a mean weight of 3.339 kg and a standard deviation of 0.573 kg.

a. What is the probability a newborn weighs more than 5 kg?

b. What is the probability a newborn weighs between 3 kg and 4 kg?

c. What fraction of newborns will be > 1.5 sd from the mean?

d. What fraction of newborns will be > 1.5 kg from the mean?

e. A random sample of 10 newborns is taken, what is the probability that their mean weight would be > 3.5 kg?

Practice Problem #6: Singleton babies born have a mean weight of 3.339 kg and a standard deviation of 0.573 kg.

a. What is the probability a newborn weighs more than 5 kg?

$$Z = \frac{X - \mu}{\sigma} = \frac{5 - 3.339}{0.573} = 2.90$$

P[Z > 2.90]=0.00187

	x.x0	x.x1	x.x2	.x3	x.x4	x.x5	x.x6
2.5	0.00621	0.00604	0.00587	0.0057	0.00554	0.00539	0.00523
2.6	0.00466	0.00453	0.0044	0.00427	0.00415	0.00402	0.00391
2.7	0.00347	0.00336	0.00326	0.00317	0.00307	0.00298	0.00289
2.8	0.00256	0.00248	0.0024	0.00233	0.00226	0.00219	0.00212
2.9	0.00187	0.00181	0.00175	0.00169	0.00164	0.00159	0.00154

Practice Problem #6: Singleton babies born have a mean weight of 3.339 kg and a standard deviation of 0.573 kg.

a. What is the probability a newborn weighs more than 5 kg?
P[Z>2.90]=0.00187

b. What is the probability a newborn weighs between 3 kg and 4 kg?

$$Z = \frac{X-\mu}{\sigma} = \frac{3-3.339}{0.573} = -0.59$$

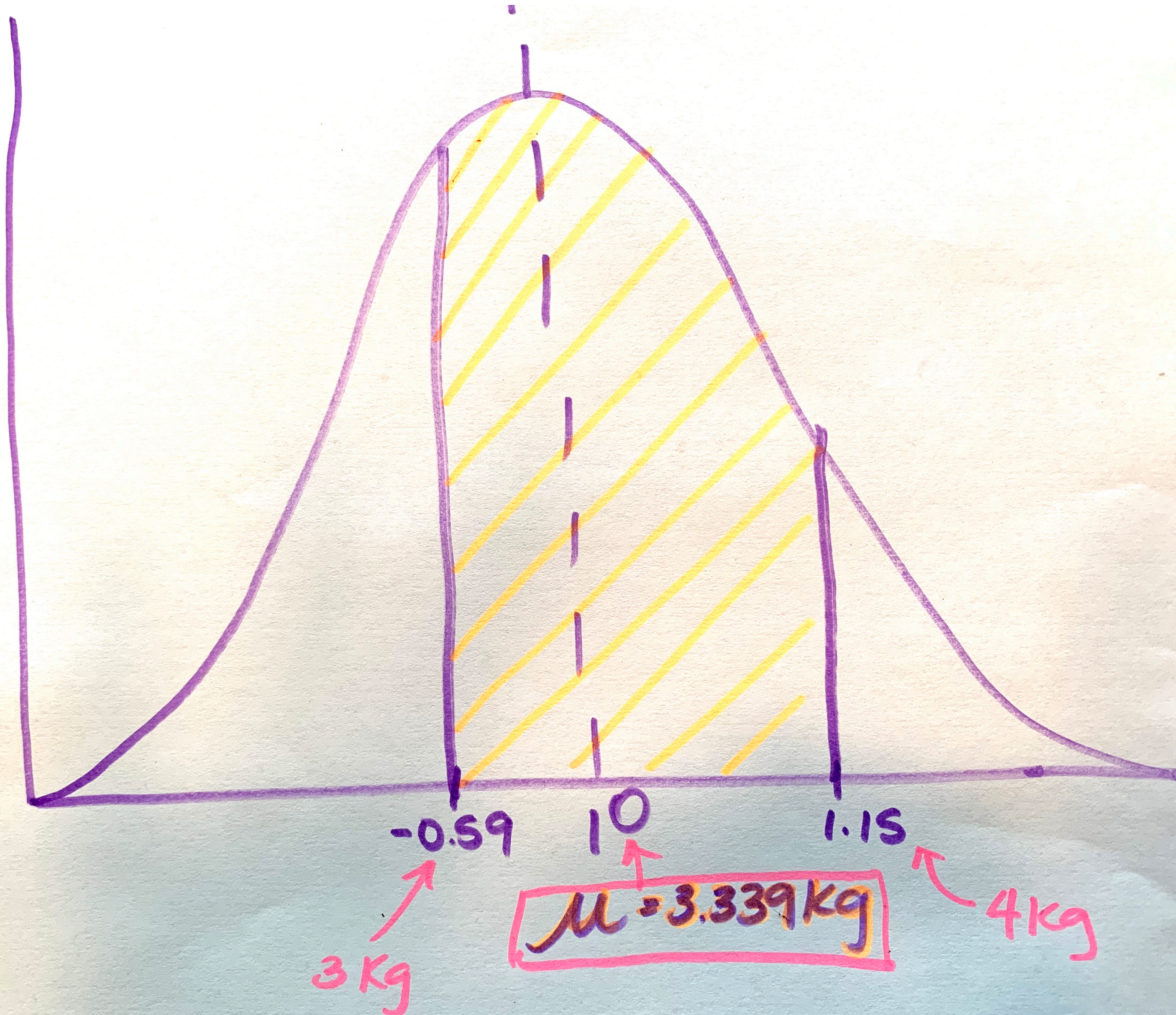
$$Z = \frac{X-\mu}{\sigma} = \frac{4-3.339}{0.573} = 1.15$$

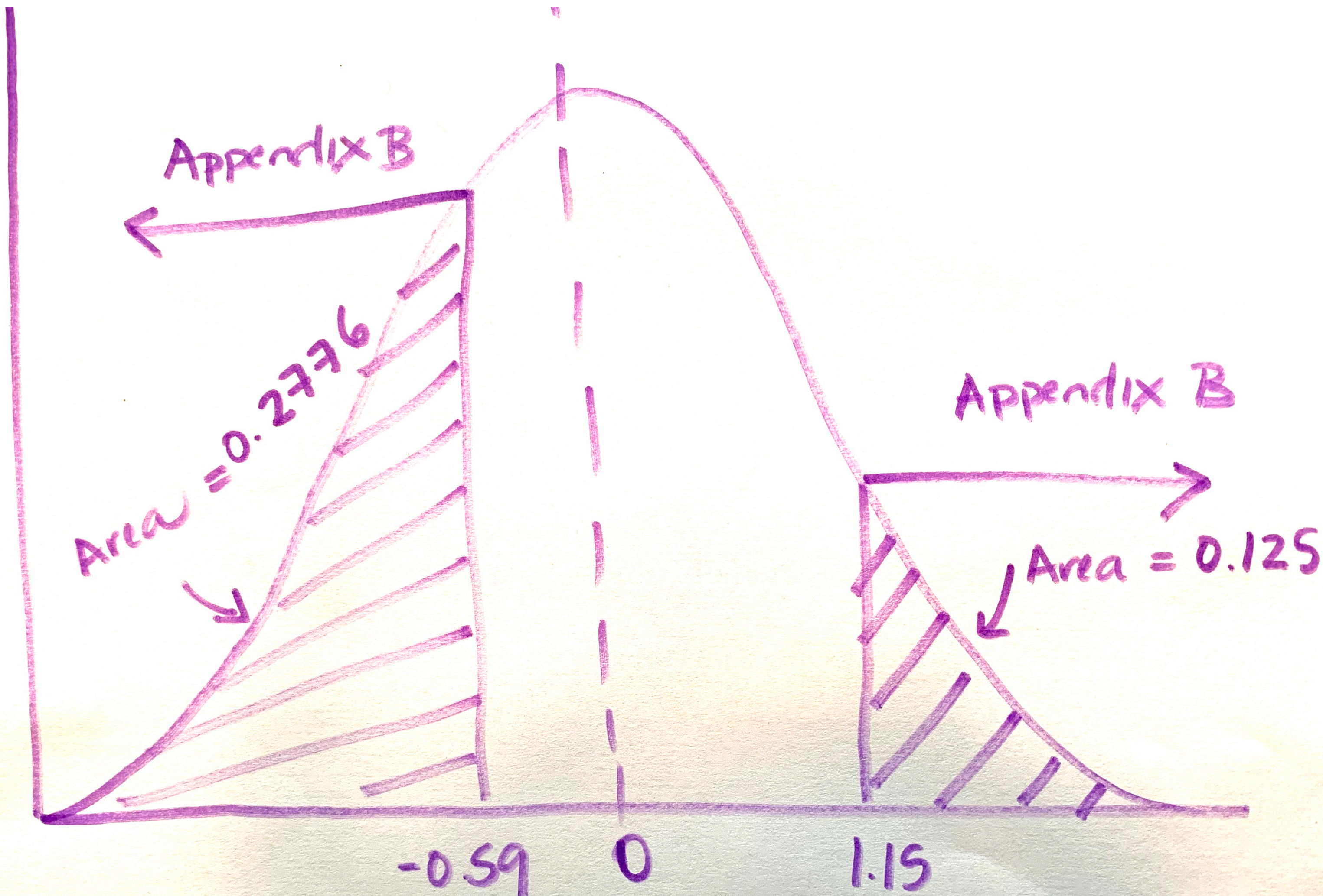
P[Z < -0.59] = P[Z > 0.59] = 0.2776

P[Z > 1.15] = 0.125

Area between these two boundaries:

1 – (0.2776+0.125) = 1 – (0.4026) = 0.5974


-0.59
0
1.15
3 kg
μ = 3.339 kg
4 kg



$P(\text{baby weighs b/w 3 and 4 Kg})$

$= 1 - (0.2776 + 0.125)$

Applications of the Normal Dist'n

Practice Problem #6 Singleton babies born have a mean weight of 3.339 kg and a standard deviation of 0.573 kg.

a. What is the probability a newborn weighs more than 5 kg?
P[Z>2.90]=0.00187

b. What is the probability a newborn weighs between 3 kg and 4 kg? 0.5974

c. What fraction of newborns will be > 1.5 sd from the mean?

- Table B gives 0.06681 of babies are ≥ 1.5 sd

1.2	0.11507	0.11314	0.11123	0.10935	0.10749	0.10565
1.3	0.09680	0.09510	0.09342	0.09176	0.09012	0.08851
1.4	0.08076	0.07927	0.07780	0.07636	0.07493	0.07353
1.5	$\boxed{0.06681}$	0.06552	0.06426	0.06301	0.06178	0.06057
1.6	0.05480	0.05370	0.05262	0.05155	0.05050	0.04947

- Symmetrical distribution so 0.134 babies are ≥ 1.5 sd or ≤ -1.5 sd

Applications of the Normal Dist'n

Practice Problem #6: Singleton babies born have a mean weight of 3.339 kg and a standard deviation of 0.573 kg.

a. What is the probability a newborn weighs more than 5 kg?
P[Z>2.90]=0.00187

b. What is the probability a newborn weighs between 3 kg and 4 kg? 0.59733

c. What fraction of newborns will be > 1.5 sd from the mean? 0.134

d. What fraction of newborns will be > 1.5 kg from the mean?

– transform the 1.5 kg into normal standard deviations

$$\frac{1.5\text{ kg}}{0.573\text{ kg}} = 2.62$$

P[Z > 2.62] = 0.0044

Once again: direction wasn't specified so… double it! 0.0088

Practice Problem #6: Singleton babies born have a mean weight of 3.339 kg and a standard deviation of 0.573 kg.

a. What is the probability a newborn weighs more than 5 kg? P[Z>2.90]=0.00187

b. What is the probability a newborn weighs between 3 kg and 4 kg? 0.59733

c. What fraction of newborns will be > 1.5 sd from the mean? 0.134

d. What fraction of newborns will be > 1.5 kg from the mean? 0.0088

e. A random sample of 10 newborns is taken, what is the probability that their mean weight would be > 3.5 kg?

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{3.5 - 3.339}{\frac{0.573}{\sqrt{10}}} = \frac{3.5 - 3.339}{0.18} = 0.89$$

- The means of samples taken from a normal distribution are themselves normally distributed but….

- ….the sampling distribution of sample means is *approximately normal* **even when the distribution of Y is <u>not normal</u>** *if the sample size is large enough*

Central Limit Theorem:

The sum (or mean) of a large number of measurements randomly sampled from any population is approximately normally distributed

Central Limit Theorem:

The sum or mean of a large number of measurements randomly sampled from *any* population is approximately normally distributed

Some extra points:

- What defines a "large number" depends on the shape of the data
- When you add up a lot of chance events --> normal distribution
  - Intelligence, as measured by IQ tests, is bell shaped…as are MANY biological/environmental traits
  - Slightly tricky assumption when dealing with disease literature: Normal processes involve "addition" whereas disease processes often involve multiplication (ie. Cancer)

The Central Limit theorem:

$$\boxed{n^{1/2}\frac{(\bar{X}_n-\mu)}{\sigma}\xrightarrow{n\to\infty}N(0,1)}$$

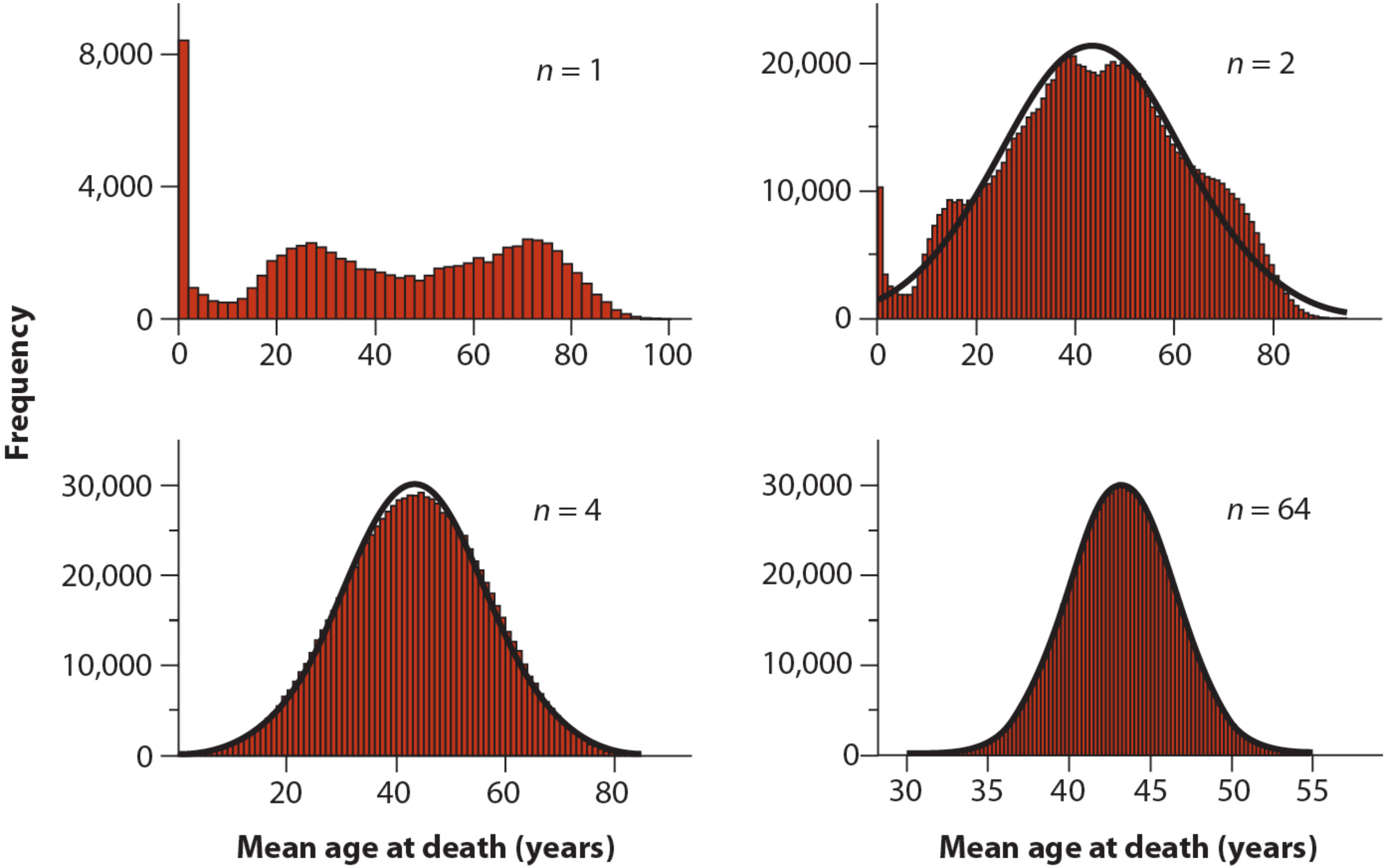
Equivalently:

$$\frac{\sum_{j=1}^{n}X_j-n\mu}{\sqrt{n}\sigma}\xrightarrow{n\to\infty}N(0,1)$$

Remember: We already know (Law of Large numbers) that:

$$\bar{X}_n-\mu\to 0$$

Applications of the Normal Dist'n



Which of the follow statements about distributions is not true?

a. The normal distribution can be used as an approximation to the binomial distribution

b.The Poisson distribution is used under conditions where n approaches 0 while p approaches infinity

c. The binomial distribution expresses the probability of getting X successes out of n trials

d. As degrees of freedom increase, the X^2 distribution approaches a normal distribution

Normal Approximation to the Binomial Distribution*:

- Remember the binomial distribution?
  - discrete
  - number of successes in n independent trials n
- Number of successes is a sum
- mean = np
- stand dev = $\sqrt{np(1-p)}$

Standard Normal Approximation to the Binomial Distribution:

1. State H_0 and H_A
2. Test Statistic
3. P-Value or Critical value/Compare to critical value (remember to double it!)

$$\boxed{P[NumSuccesses \geq X] = P[Z > \frac{X - np}{\sqrt{np(1-p)}}]}$$

4. State a conclusion