STA286 Lecture 20

Neil Montgomery

Last edited: 2017-03-09 13:01

basic result for $\overline{X}$ in general

For a sample $X_1, \ldots, X_n$ i.i.d. from *any* distribution with mean μ and variance σ^2, the following are always true:

$$
\begin{aligned}
E\left(\overline{X}\right) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
\mathrm{Var}\left(\overline{X}\right) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)
\end{aligned}
$$

basic result for $\overline{X}$ in general

For a sample $X_1, \ldots, X_n$ i.i.d. from *any* distribution with mean μ and variance σ^2, the following are always true:

$$
\begin{aligned}
E\left(\overline{X}\right) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) \\
\mathrm{Var}\left(\overline{X}\right) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)
\end{aligned}
$$

basic result for $\overline{X}$ in general

For a sample $X_1, \ldots, X_n$ i.i.d. from *any* distribution with mean μ and variance σ^2, the following are always true:

$$E\left(\overline{X}\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}n\mu = \mu$$

$$\mathrm{Var}\left(\overline{X}\right) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

basic result for $\overline{X}$ in general

For a sample $X_1, \ldots, X_n$ i.i.d. from *any* distribution with mean μ and variance σ^2, the following are always true:

$$E\left(\overline{X}\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} n\mu = \mu$$

$$\mathrm{Var}\left(\overline{X}\right) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i)$$

basic result for $\overline{X}$ in general

For a sample $X_1, \ldots, X_n$ i.i.d. from *any* distribution with mean μ and variance σ^2, the following are always true:

$$E\left(\overline{X}\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} n\mu = \mu$$

$$\mathrm{Var}\left(\overline{X}\right) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

basic result for $\overline{X}$ in general

For a sample $X_1, \ldots, X_n$ i.i.d. from *any* distribution with mean μ and variance σ^2, the following are always true:

$$
\begin{aligned}
E\left(\overline{X}\right) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} n\mu = \mu \\
\mathrm{Var}\left(\overline{X}\right) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n} \\
SD(\overline{X}) &= \frac{\sigma}{\sqrt{n}}
\end{aligned}
$$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)^n$$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)^n = e^{n\mu t + n\sigma^2 t^2/2}$$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)^n = e^{n\mu t + n\sigma^2 t^2/2}$$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)^n = e^{n\mu t + n\sigma^2 t^2/2}$$

so that $\sum_{i=1}^{n} X_i \sim N\left(n\mu, \sqrt{n}\sigma\right)$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)^n = e^{n\mu t + n\sigma^2 t^2/2}$$

so that $\sum_{i=1}^{n} X_i \sim N\left(n\mu, \sqrt{n}\sigma\right)$

Using the rules for normal distributions we also get:

$$\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)^n = e^{n\mu t + n\sigma^2 t^2/2}$$

so that $\sum_{i=1}^{n} X_i \sim N\left(n\mu, \sqrt{n}\sigma\right)$

Using the rules for normal distributions we also get:

$$\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \qquad\qquad \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum\limits_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)^n = e^{n\mu t + n\sigma^2 t^2/2}$$

so that $\sum\limits_{i=1}^{n} X_i \sim N\left(n\mu, \sqrt{n}\sigma\right)$

Using the rules for normal distributions we also get:

$$\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \qquad \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

full distribution of $\overline{X}$ when sample is normal

For a sample $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma)$ we have:

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left(e^{\mu t + \sigma^2 t^2/2}\right)^n = e^{n\mu t + n\sigma^2 t^2/2}$$

so that $\sum_{i=1}^{n} X_i \sim N\left(n\mu, \sqrt{n}\sigma\right)$

Using the rules for normal distributions we also get:

$$\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \qquad\qquad \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

The seemingly impossible task is to determine the distribution of $\overline{X}$ when the underlying distribution is not normal (i.e., almost always.)

the actual central limit theorem

For any random variable X with mean μ and variance σ^2, consider a sample $X_1, \ldots, X_n$ from the same distribution.

Now consider the sample average of this sample: $\overline{X}_n$ (because n will be changing below. . .).

the actual central limit theorem

For any random variable X with mean μ and variance σ^2, consider a sample $X_1, \ldots, X_n$ from the same distribution.

Now consider the sample average of this sample: $\overline{X}_n$ (because n will be changing below. . .).

The actual Central Limit Theorem (CLT) says:

$$P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leqslant u\right)$$

the actual central limit theorem

For any random variable X with mean μ and variance σ^2, consider a sample $X_1, \ldots, X_n$ from the same distribution.

Now consider the sample average of this sample: $\overline{X}_n$ (because n will be changing below...).

The actual Central Limit Theorem (CLT) says:

$$\lim_{n \to \infty} P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leqslant u\right)$$

the actual central limit theorem

For any random variable X with mean μ and variance σ^2, consider a sample $X_1, \ldots, X_n$ from the same distribution.

Now consider the sample average of this sample: $\overline{X}_n$ (because n will be changing below. . .).

The actual Central Limit Theorem (CLT) says:

$$\lim_{n\to\infty} P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leqslant u\right) = P(Z \leqslant u)$$

where $Z \sim N(0, 1)$.

the actual central limit theorem

For any random variable X with mean μ and variance σ^2, consider a sample $X_1, \ldots, X_n$ from the same distribution.

Now consider the sample average of this sample: $\overline{X}_n$ (because n will be changing below...).

The actual Central Limit Theorem (CLT) says:

$$\lim_{n\to\infty} P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leqslant u\right) = P(Z \leqslant u)$$

$$\lim_{n\to\infty} F_n(u) = F_Z(u)$$

where $Z \sim N(0, 1)$.

the value of the CLT is in the speed of convergence

Useful limit theorems are ones where the convergence is fast—the CLT is such an example.

$$P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leqslant u\right) \approx P(Z \leqslant u)$$

for n "large enough".

the value of the CLT is in the speed of convergence

Useful limit theorems are ones where the convergence is fast—the CLT is such an example.

$$P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leqslant u\right) \approx P(Z \leqslant u)$$

for n "large enough".

How large? Depends on the shape of the underlying distribution X.

- $n = 2$ would need X normal.

the value of the CLT is in the speed of convergence

Useful limit theorems are ones where the convergence is fast—the CLT is such an example.

$$P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leqslant u\right) \approx P(Z \leqslant u)$$

for n "large enough".

How large? Depends on the shape of the underlying distribution X.

- $n = 2$ would need X normal.
- $n = 10$ good enough for symmetric distributions without outliers.

the value of the CLT is in the speed of convergence

Useful limit theorems are ones where the convergence is fast—the CLT is such an example.

$$P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leqslant u\right) \approx P(Z \leqslant u)$$

for n "large enough".

How large? Depends on the shape of the underlying distribution X.

- $n = 2$ would need X normal.
- $n = 10$ good enough for symmetric distributions without outliers.
- $n = 30$ for mildly skewed X.

the value of the CLT is in the speed of convergence

Useful limit theorems are ones where the convergence is fast—the CLT is such an example.

$$P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leqslant u\right) \approx P(Z \leqslant u)$$

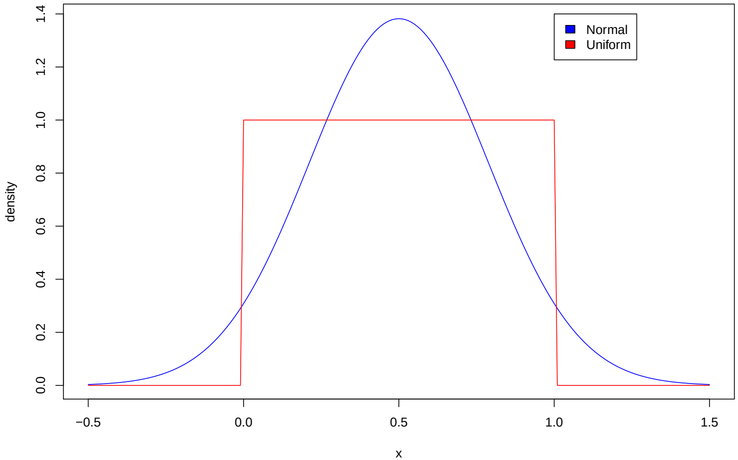
for n "large enough".

How large? Depends on the shape of the underlying distribution X.

- $n = 2$ would need X normal.
- $n = 10$ good enough for symmetric distributions without outliers.
- $n = 30$ for mildly skewed X.
- $n > 60$ for more skewed.

how large is large enough case I - symmetric/no outliers

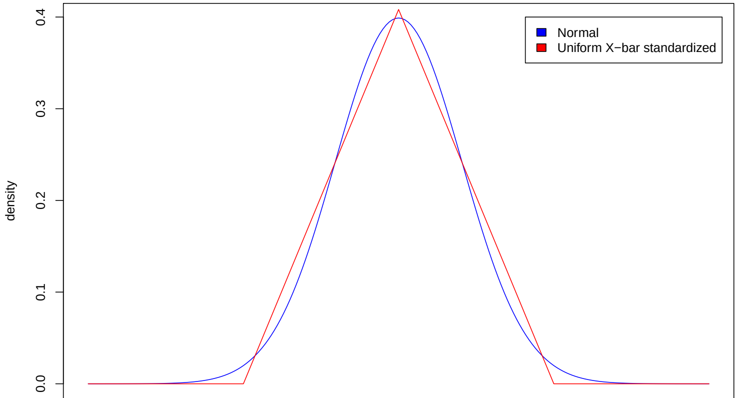
Consider a Uniform[0,1] distribution. The mean is 0.5 and the standard deviation is $1/\sqrt{12}$. Here is a plot of a uniform density versus a $N(0.5, 1/\sqrt{12})$ density:



how large is large enough case I - symmetric/no outliers

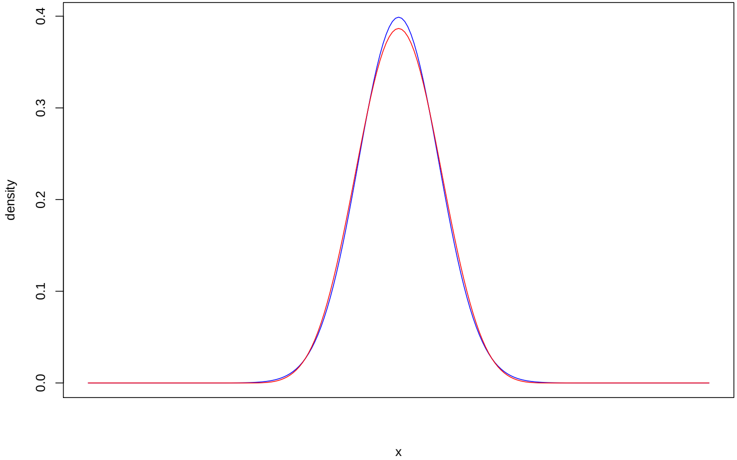
Now consider X_1, X_2 i.i.d. Uniform[0,1], and its sample average $\overline{X}$, which will have mean 0.5 and standard deviation $1/\sqrt{12 \cdot 2}$.

Here is a picture of the density for $\frac{\overline{X}-0.5}{(1/\sqrt{12 \cdot 2})/\sqrt{2}}$, along with the density for $Z \sim N(0,1)$.



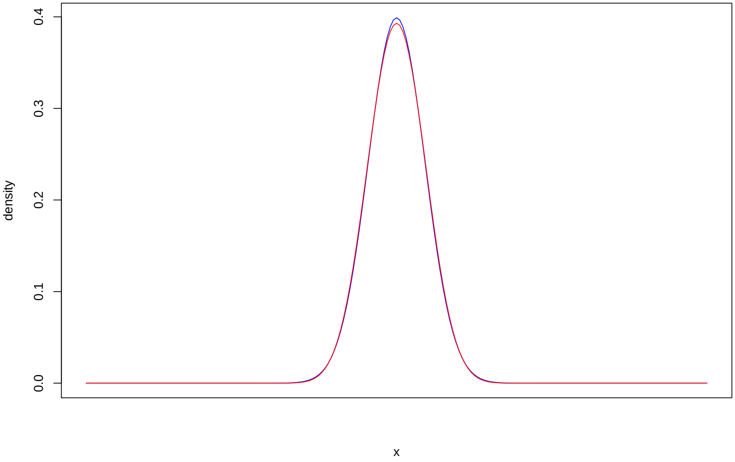
how large is large enough case I - symmetric/no outliers

Same as before but now with X_1, X_2, X_3, X_4, X_5 (i.e. $n = 5$):



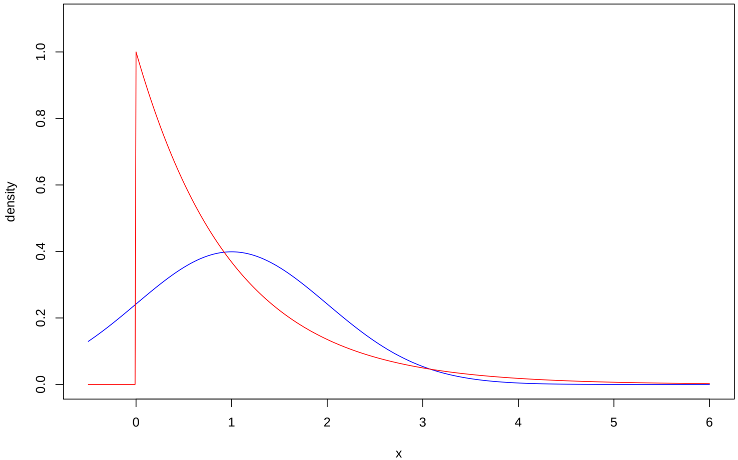
how large is large enough case I - symmetric/no outliers

Now with $n = 10$.



how large is large enough case II - skewed

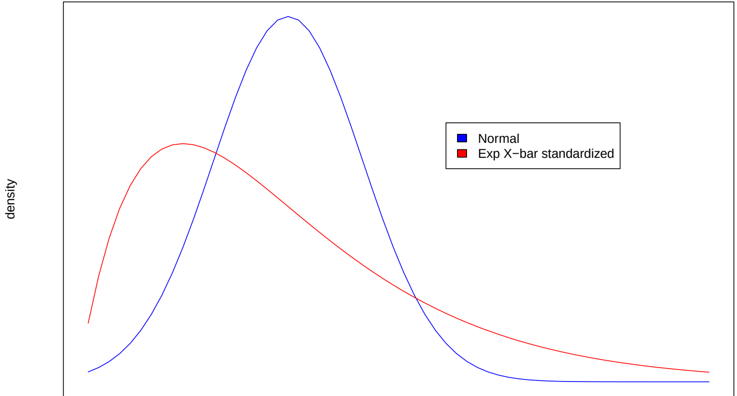
Now let the "underlying" distribution be Exp(1), which has mean and standard deviation both equal to $1/1 = 1$. Here's the density along with a $N(1, 1)$ density:



how large is large enough case II - skewed

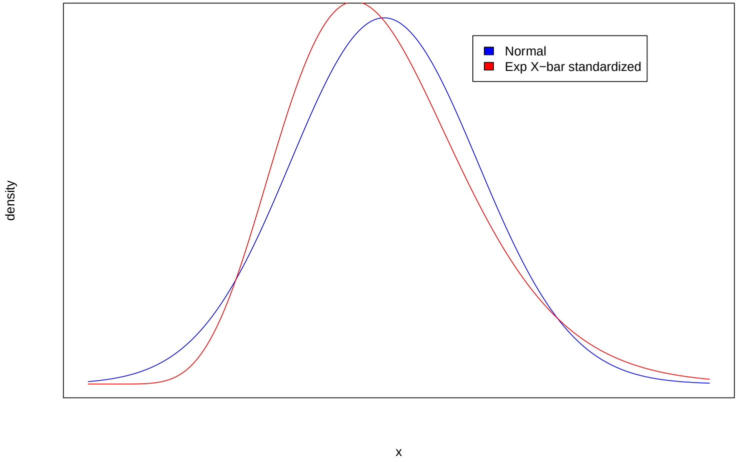
Now consider X_1, X_2 i.i.d. Exp(1), and its sample average $\overline{X}$, which will have mean 1 and standard deviation $1/\sqrt{2}$.

Here is a picture of the density for $\frac{\overline{X}-1}{1/\sqrt{2}}$, along with the density for $Z \sim N(0,1)$.



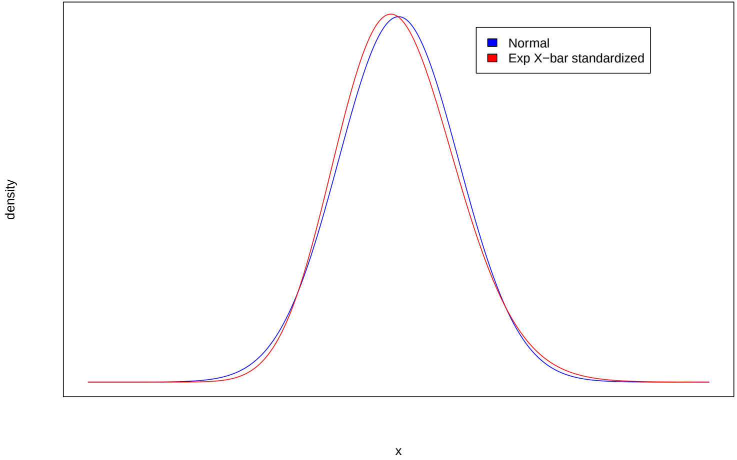
how large is large enough case II - skewed

Now with $n = 10$.



how large is large enough case II - skewed

Try $n = 60$



normal approximation applications

Given $X_1, \ldots, X_n$ i.i.d with mean μ and variance σ^2, as long as n is large enough, any of the following approximations hold. Pick the most convenient:

$$\sum_{i=1}^{n} X_i \sim^{approx} N(n\mu, \sqrt{n}\sigma)$$

$$\overline{X} \sim^{approx} N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim^{approx} N(0, 1)$$

normal approximation example

Piston lifetimes in a Diesel engine follow a roughly symmetric distribution with mean 3.4 years and standard deviation 2.1 years.

What is the chance that the average life of 25 engines exceeds 4 years?

$$P(\overline{X} > 4)$$

normal approximation example

Piston lifetimes in a Diesel engine follow a roughly symmetric distribution with mean 3.4 years and standard deviation 2.1 years.

What is the chance that the average life of 25 engines exceeds 4 years?

$$P(\overline{X} > 4) = P\left(\frac{\overline{X} - 3.4}{2.1/\sqrt{25}} > \frac{4 - 3.4}{2.1/\sqrt{25}}\right)$$

normal approximation example

Piston lifetimes in a Diesel engine follow a roughly symmetric distribution with mean 3.4 years and standard deviation 2.1 years.

What is the chance that the average life of 25 engines exceeds 4 years?

$$P(\overline{X} > 4) = P\left(\frac{\overline{X} - 3.4}{2.1/\sqrt{25}} > \frac{4 - 3.4}{2.1/\sqrt{25}}\right) \approx P(Z > 1.43)$$

normal approximation example

Piston lifetimes in a Diesel engine follow a roughly symmetric distribution with mean 3.4 years and standard deviation 2.1 years.

What is the chance that the average life of 25 engines exceeds 4 years?

$$P(\overline{X} > 4) = P\left(\frac{\overline{X} - 3.4}{2.1/\sqrt{25}} > \frac{4 - 3.4}{2.1/\sqrt{25}}\right) \approx P(Z > 1.43) = 0.0763585$$

another normal approximation example

A defective item is produced with probability $p = 0.01$. After $n = 10000$ items are produced, what is the probability that there were fewer than 80 defective items produced?

another normal approximation example

A defective item is produced with probability $p = 0.01$. After $n = 10000$ items are produced, what is the probability that there were fewer than 80 defective items produced?

If X_i is 0 or 1 as item i is not defective, or defective, respectively, then $X_i \sim \text{Bernoulli}(0.01)$. We want $P\left(\sum_{i=1}^{n} X_i < 80\right)$.

another normal approximation example

A defective item is produced with probability $p = 0.01$. After $n = 10000$ items are produced, what is the probability that there were fewer than 80 defective items produced?

If X_i is 0 or 1 as item i is not defective, or defective, respectively, then $X_i \sim \text{Bernoulli}(0.01)$. We want $P\left(\sum_{i=1}^{n} X_i < 80\right)$.

In principle this is a $\text{Binomial}(n, p)$ calculation, but a very difficult one.