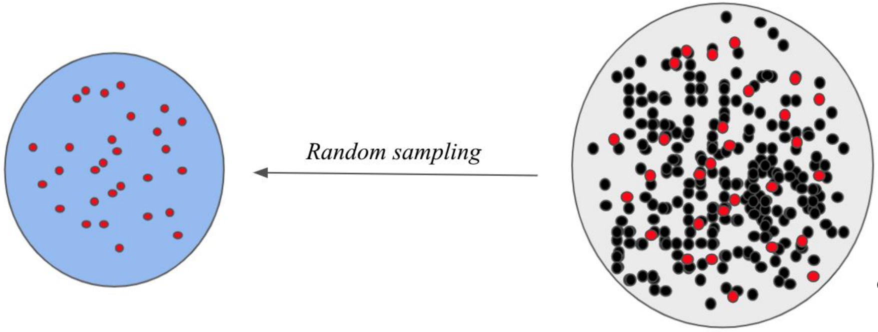
Estimate population parameters

under the i.i.d assumption

06/22/2021

Berkeley
UNIVERSITY OF CALIFORNIA

In the previous lecture,



- Population vs. samples
  - Finite population: simple random sampling
    Cannot assume i.i.d. (Not required in this course)
  - Finite population: sampling **with** replacement
    Can assume i.i.d.
  - Infinite population:
    Can assume i.i.d.

- Estimator $\hat{\theta}_n = g(X_1, \ldots, X_n)$ for popular parameter θ
  Sample mean $\bar{X}_n$ *was used as an estimator of* μ.

- Unbiased estimators:

$$\text{bias}(\hat{\theta}_n) = E\hat{\theta}_n - \theta = 0.$$

Under i.i.d, sample mean $\bar{X}_n$ *is an* *<u>unbiased estimator</u>* *of* μ .

How to evaluate an estimator?

An estimator of a model parameter θ is a function of the sample $X_1, \ldots, X_n$.

Definition. We evaluate the 'goodness' of an estimator $\hat{\theta}_n$ by the *mean squared error* (MSE):

$$\mathrm{MSE}(\hat{\theta}_n) = E\left(\hat{\theta}_n - \theta\right)^2.$$

$$\mathrm{MSE}(\hat{\theta}_n) = \underbrace{\mathrm{Var}\left(\hat{\theta}_n\right)}_{\text{variance}} + \underbrace{\left[E\left(\hat{\theta}_n\right) - \theta\right]^2}_{\text{bias}}$$

Proof. On page 19 of this set of slides, we showed for any random variable

$$\mathrm{Var}(X) = E\left(X^2\right) - [E(X)]^2 \qquad (2)$$

Apply (2) to the expectation in the definition to get

$$\mathrm{MSE}(\hat{\theta}_n) = E\left(\hat{\theta}_n^2 + \theta^2 - 2\hat{\theta}_n\theta\right) = \underbrace{E\left(\hat{\theta}_n^2\right)}_{\text{use (2)}} + \theta^2 - 2\theta E\left(\hat{\theta}_n\right)$$

$$= \mathrm{Var}(\hat{\theta}_n) + \underbrace{\left[E\left(\hat{\theta}_n\right)\right]^2 + \theta^2 - 2\theta E\left(\hat{\theta}_n\right)}_{=\left[\theta - E\left(\hat{\theta}_n\right)\right]^2}.$$

$\square$

Estimate population variance

Population variance: $\sigma^2 = \int_{\chi} (x-\mu)^2 f(x)dx$ $\longleftrightarrow$ Sample variance: $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$

Theorem B. Under the i.i.d assumption, $E\left(\hat{\sigma}_n^2\right) = \frac{n-1}{n}\sigma^2$.

*Is $\hat{\sigma}_n^2$ an **unbiased estimator** of σ^2?*

Estimate population parameters

What if we adjust slightly: $\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$

Theorem B'. Under the i.i.d assumption, $E\left(\hat{\sigma}_n^2\right) = \sigma^2$.

*Is $\hat{\sigma}_n^2$ an **unbiased estimator** of σ^2 now?*

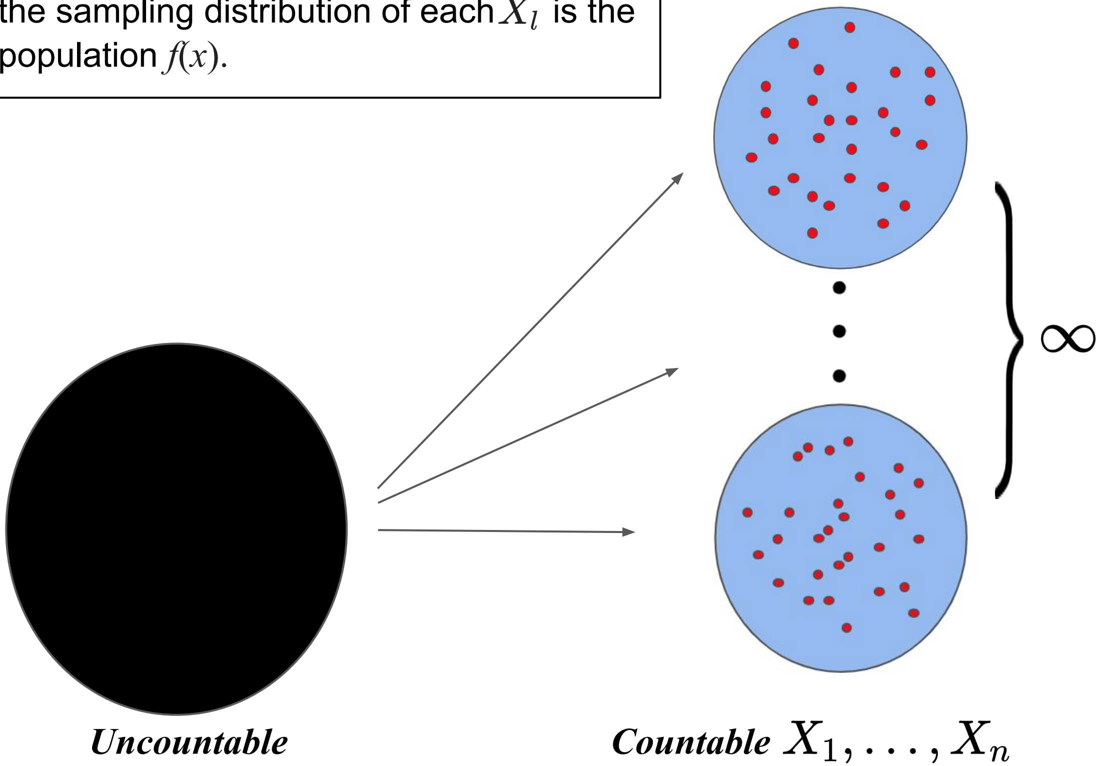
Proof is left as an exercise in HW1.

Normal approximation of $\bar{X}_n$

06/23/2021

Under the i.i.d setting

When the probability space is continuous, the sampling distribution of each X_l is the population $f(x)$.



What about the sampling distribution of $\bar{X}_n$*?*

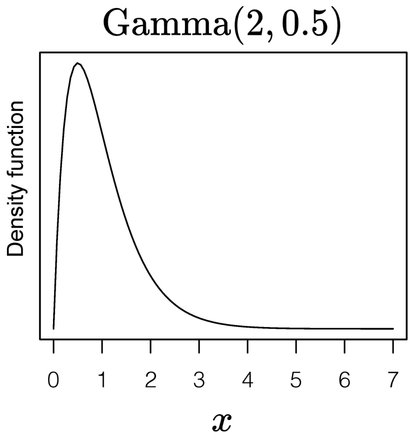
$$E(\bar{X}_n) = \mu$$

$$\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$$

Simulation:
Draw as many samples of size n as possible, and plot the histogram of all $\bar{X}_n$.

Under the i.i.d setting - $n = 2$

Suppose $X_1, \ldots, X_n$ are i.i.d samples from Gamma(2, 0.5).



In this case, $\mu = 2/0.5 = 4,\ \sigma^2 = 8$.

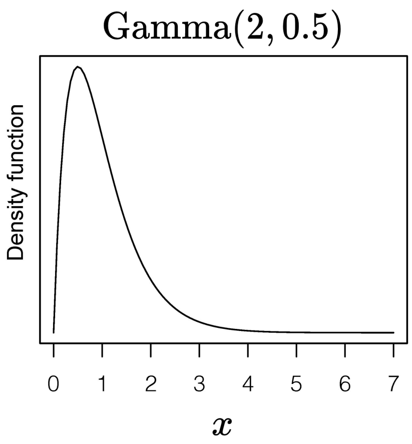
$$E(\bar{X}_n) = 4$$
$$\mathrm{Var}(\bar{X}_n) = \frac{8}{2} = 4$$

Actually, we know exactly $\dfrac{X_1 + X_2}{2} \sim \mathrm{Gamma}(4, 1)$.

Simulation:
Draw as many samples of size n as possible, and plot the histogram of all $\bar{X}_n$.

Under the i.i.d setting - $n = 2$

Suppose $X_1, \ldots, X_n$ are i.i.d samples from Gamma(2, 0.5).



In this case, $\mu = 2/0.5 = 4,\ \sigma^2 = 8$.

$$E(\bar{X}_n) = 4$$

$$\mathrm{Var}(\bar{X}_n) = \frac{8}{2} = 4$$

Actually, we know exactly $\dfrac{X_1 + X_2}{2} \sim \mathrm{Gamma}(4,\ 1)$.

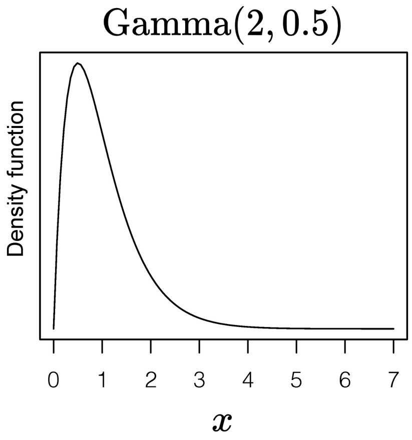
The number of samples $\nearrow$ $\Rightarrow$
The accuracy of the histogram $\nearrow$

Simulation:
Draw as many samples of size n as possible, and plot the histogram of all $\bar{X}_n$.

Under the i.i.d setting - $n = 100$

Suppose $X_1, \ldots, X_n$ are i.i.d samples from Gamma(2, 0.5).

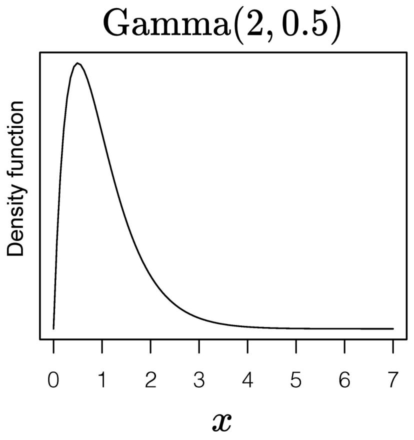
We still know exactly $\dfrac{X_1 + \ldots + X_{100}}{100} \sim \text{Gamma}(200,\, 100)$

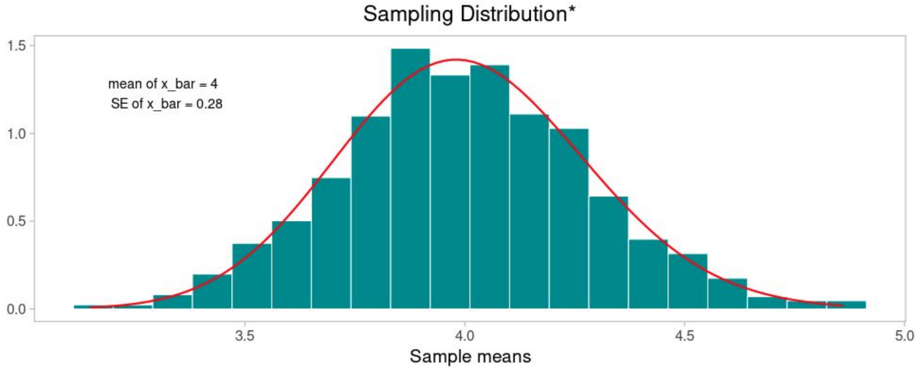


Under the i.i.d setting - $n = 100$

Suppose $X_1, \ldots, X_n$ are i.i.d samples from Gamma(2, 0.5).

We still know exactly $\dfrac{X_1 + \ldots + X_{100}}{100} \sim \text{Gamma}(200,\, 100)$





* The histogram is showing 900 samples with each sample containing n=100 observations. The red curve is the density plot of Gamma(200,100).

The distribution of $\bar{X}_n$ *became symmetric!*

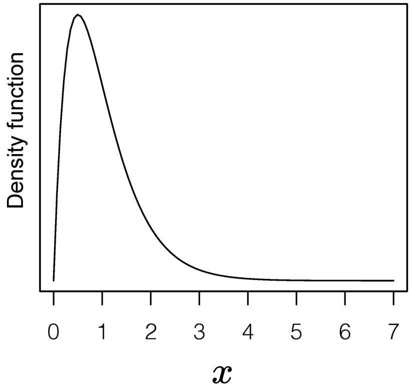
Under the i.i.d setting - $n = 100$

$X \sim \text{Beta}$

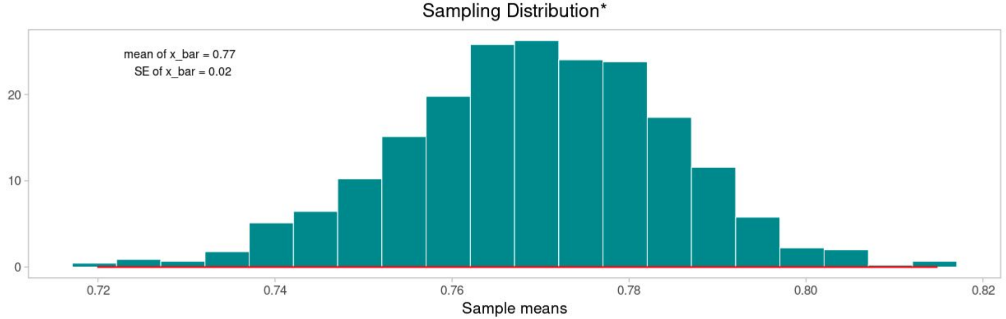
$Y \sim \text{Beta}$ $\quad X+Y$

Suppose $X_1, \ldots, X_n$ are i.i.d samples from other distributions.

$\text{Gamma}(2, 0.5)$



It is hard to compute the distribution of $\dfrac{X_1 + \ldots + X_{100}}{100}$.



* The histogram is showing 900 samples with each sample containing n=100 observations of a left-skewed distribution. We <u>don't know</u> the exact distribution of the sample mean.

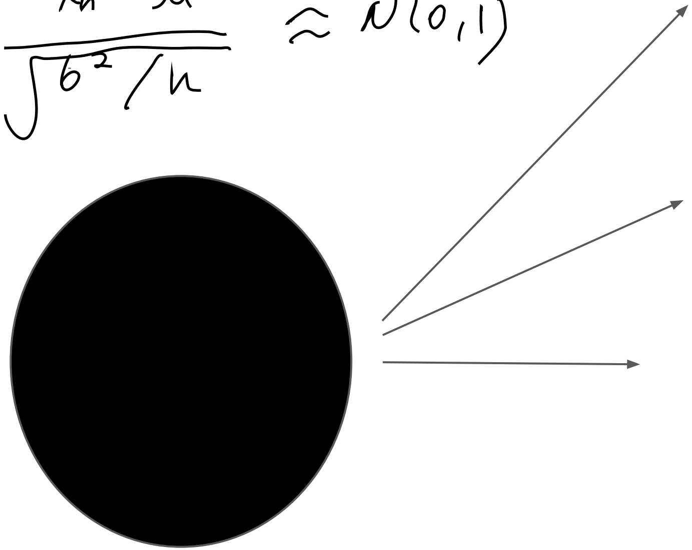
The distribution of $\bar{X}_n$ *is still symmetric!*

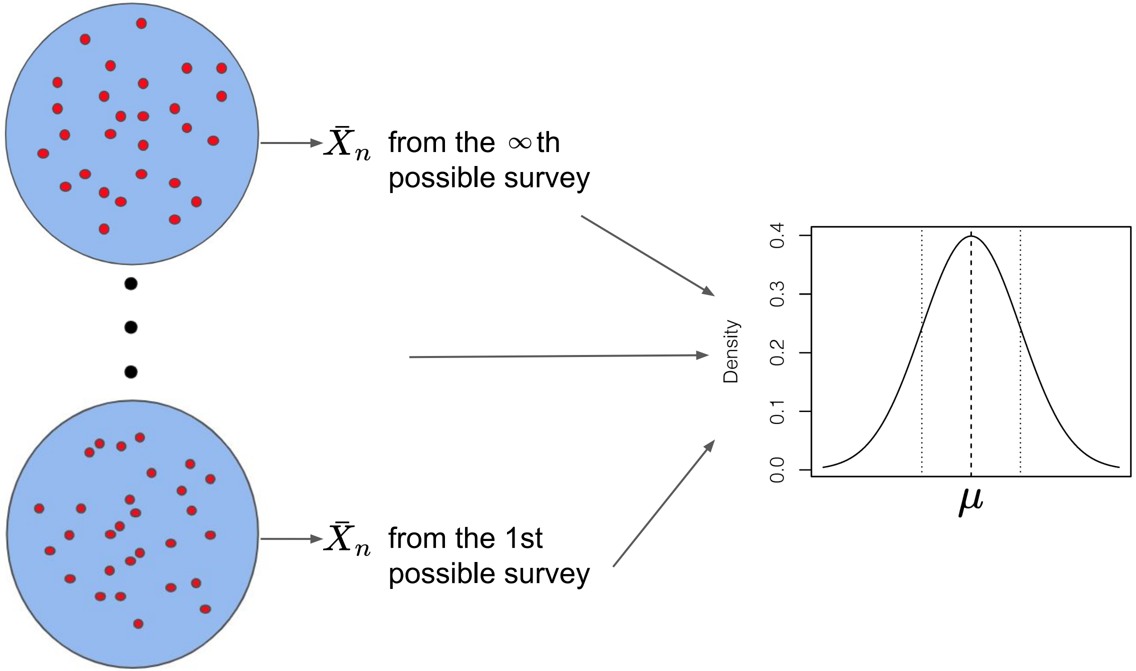
Central limit theorem

Under the i.i.d assumption,

$$\bar{X}_n \to N\left(\mu, \frac{\sigma^2}{n}\right), \text{ as } n \to \infty$$

$$\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2 / n}} \approx N(0,1)$$





Uncountable

Countable $X_1, \ldots, X_n$

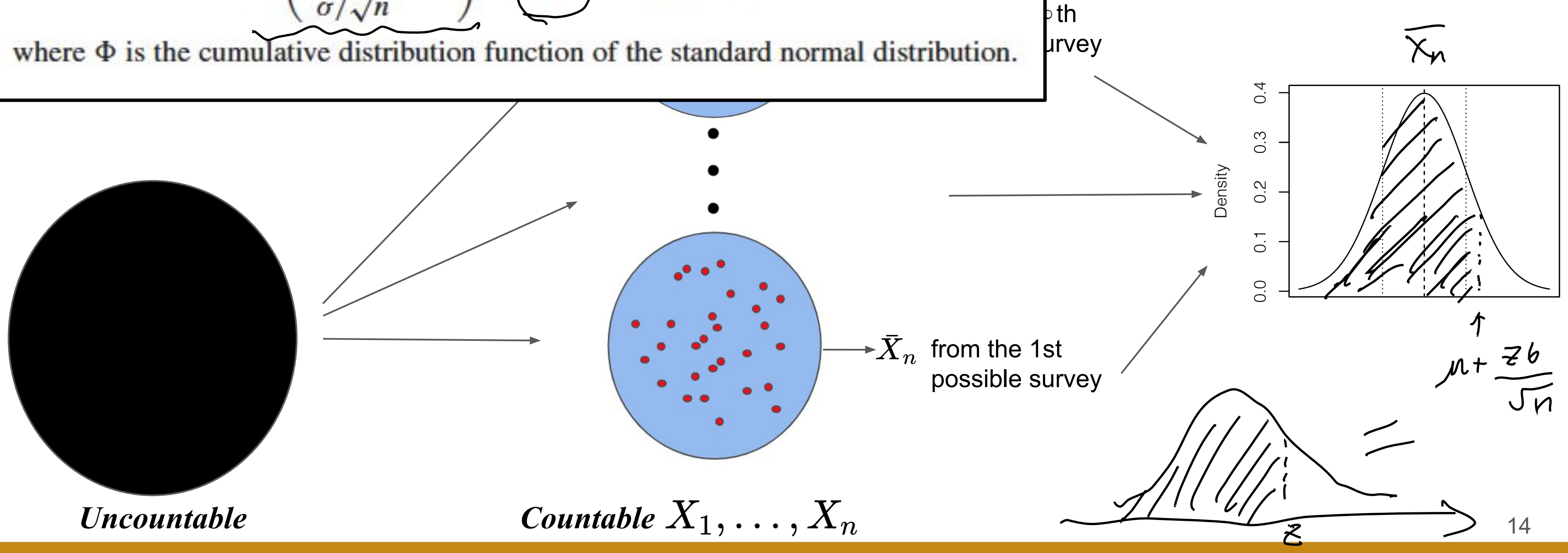
Central limit theorem - rigorous statement

Regardless of the true $f(x)$, as long as $X_1, \ldots, X_n$ are i.i.d,

$$= P\left(\bar{X}_n \le \mu + \frac{z\sigma}{\sqrt{n}}\right)$$

$$P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \le z\right) \to \Phi(z) \qquad \text{as } n \to \infty$$

where Φ is the cumulative distribution function of the standard normal distribution.



$\Phi(-x) = 1 - \Phi(x)$

Central limit theorem - application

Theorem C. Under the i.i.d assumption,

$$P(|\bar{X}_n - \mu| \le \delta) \approx 2\Phi\left(\frac{\sqrt{n}\delta}{\sigma}\right) - 1$$

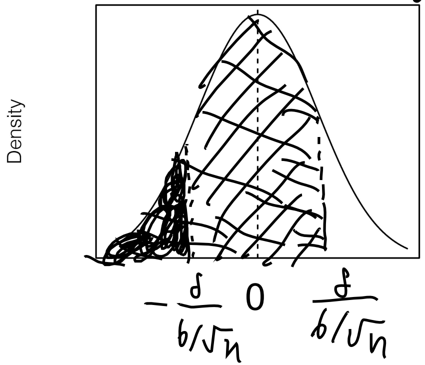
Proof. By central limit theorem, $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$.

$$P(|\bar{X}_n - \mu| \le \delta) = P\left(\frac{|\bar{X}_n - \mu|}{\sigma/\sqrt{n}} \le \frac{\delta}{\sigma/\sqrt{n}}\right)$$

$Z \sim N(0,1)$

$$\approx P\left(|Z| \le \frac{\delta}{\sigma/\sqrt{n}}\right)$$

Standard Normal



$$\Phi\left(\frac{\delta}{\sigma/\sqrt{n}}\right) - \underbrace{\Phi\left(-\frac{\delta}{\sigma/\sqrt{n}}\right)}_{=1-\Phi\left(\frac{\delta}{\sigma/\sqrt{n}}\right)}$$

$$= 2\Phi\left(\frac{\sqrt{n}\delta}{\sigma}\right) - 1. \quad \square$$

Central limit theorem - application

$X \sim \text{Bernoulli}(p)$,

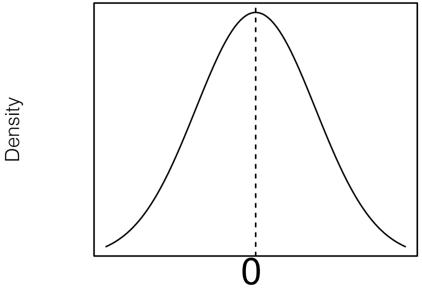
	1	0
probability	p	$1-p$

Example 1. Suppose $X_1, \ldots, X_{100}$ are i.i.d samples from Bernoulli(0.3). Approximate the following probability

$$P\left(\left|\bar{X}_{100} - \mu\right| \leq 2\right)$$

Solution: $\mu = 1 \times p + 0 \times (1-p) = p$

$$\sigma^2 = EX^2 - (EX)^2 = p - p^2 = p(1-p).$$



$$P(|\bar{X}_{100} - \mu| \leq 2) \approx 2\Phi\left(\frac{2}{\sigma/\sqrt{n}}\right) - 1$$

$$= 2\Phi\left(\frac{2}{\sqrt{p(1-p)}/\sqrt{100}}\right) - 1$$

pnorm

$\uparrow$

$$= 2\Phi\left(\frac{2}{\sqrt{0.3 \times (1-0.3)/100}}\right) - 1 = 2\Phi(43.64) - 1.$$

Confidence interval (CI) for μ

06/23/2021

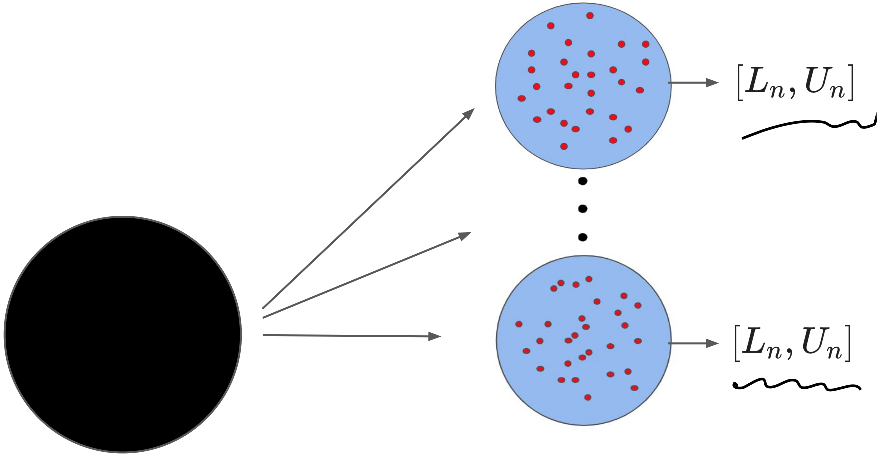
Confidence interval

$(1-\alpha)\times 100\%$ CI for θ

Definition. A confidence interval $[L_n, U_n]$ for θ, is a random interval calculated from the sample, that contains θ with some specified probability

$$P(L_n \leq \theta \leq U_n) \geq 1-\alpha.$$

Note: $L_n = g(X_1, \ldots, X_n)$, $U_n = h(X_1, \ldots, X_n)$ only depend on the samples.



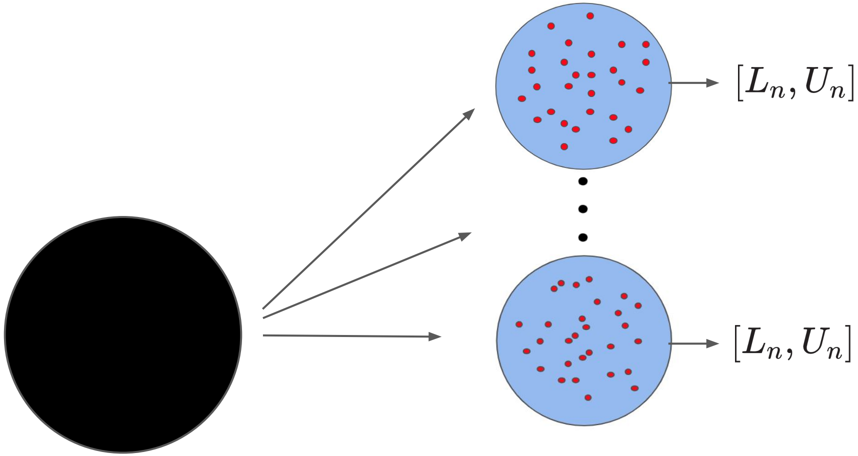
The confidence levels represents theoretical long-run frequency (i.e., the proportion) of confidence intervals that contain the true θ.

Confidence interval

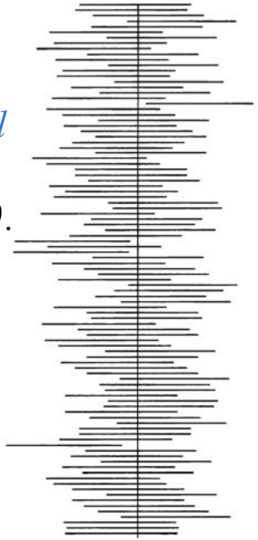
Definition. A confidence interval $[L_n, U_n]$ for θ, is a random interval calculated from the sample, that contains θ with some specified probability

$$P(L_n \leq \theta \leq U_n) \geq 1 - \alpha.$$

Note: $L_n = g(X_1, \ldots, X_n)$, $U_n = h(X_1, \ldots, X_n)$ only depend on the samples.



The confidence levels represents theoretical long-run frequency (i.e., the proportion) of confidence intervals that contain the true θ.



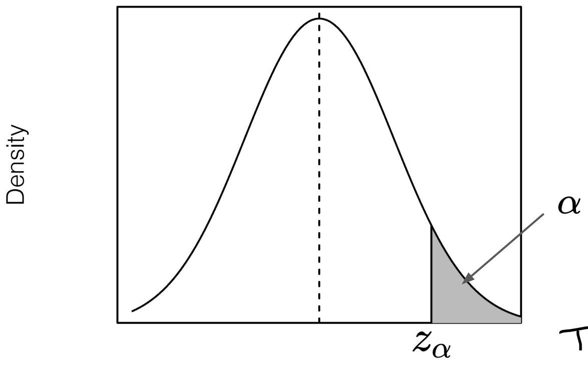
For example, if $1 - \alpha = 0.99$, 99% of **all** intervals should contain the parameter.

Confidence interval for μ

Let z_α be the number such that $\Phi(z_\alpha) = 1 - \alpha$.

Theorem D. Under the i.i.d assumption,

$$P\left(\bar{X}_n - \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \le \mu \le \bar{X}_n + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right) \approx 1 - \alpha.$$



$1-\alpha = 0.99,\ z_{\alpha/2} = 2.576$

$1-\alpha = 0.95,\ z_{\alpha/2} = 1.96$

$1-\alpha = 0.90,\ z_{\alpha/2} = 1.645$

Proof. $P\left(\bar{X}_n - \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \le \mu \le \frac{\bar{X}_n + z_{\alpha/2}\sigma}{\sqrt{n}}\right)$

$= P\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \le z_{\alpha/2} \text{ and } \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \ge -z_{\alpha/2}\right)$

$= P\left(\left|\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right| \le z_{\alpha/2}\right)$

Theorem C

$= 2\Phi(z_{\alpha/2}) - 1 = 2\left(1 - \frac{\alpha}{2}\right) - 1$

$= 1 - \alpha.$ $\square$

$-\infty \longrightarrow \infty$ $\quad P(Z = a) = 0$ $\quad 1-\alpha = 0.99$

Confidence interval for μ - Conservative estimate

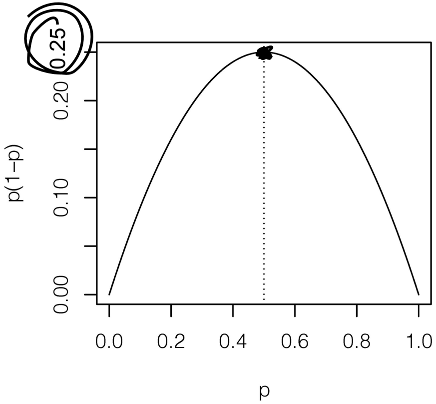
Example 2. A coin lands heads with probability p. It is tossed 100 times, and ~~4~~31 heads was observed. Compute the 99% confidence interval for p using Theorem D.

Solution. Let X_i denote the outcome of each toss, $i = 1, \dots, 100$.

We can assume i.i.d for $X_1, \dots, X_{100}$.

$X_i \sim \text{Bernoulli}(p)$, $i = 1, \dots, 100$.

$\mu = p$, $\sigma^2 = p(1-p)$. $\quad \bar{X}_n = \frac{1}{n}\sum_{i=1}^{100} X_i = \frac{41}{100} = 0.31$.



The 99% confidence interval for p will be:

$$\left[\bar{X}_n - \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}},\ \bar{X}_n + \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}}\right]$$

$$\subseteq \left[\bar{X}_n - \frac{z_{\alpha/2}\cdot 0.5}{\sqrt{n}},\ \bar{X}_n + \frac{z_{\alpha/2}\cdot 0.5}{\sqrt{n}}\right]$$

$$= 0.31 \pm \frac{2.576 \times 0.5}{\sqrt{100}} = [0.1812,\ 0.4388].$$

$\sigma^2 \le \frac{1}{4}$

$\sigma \le 0.5$

$1-\alpha = 0.99,\ z_{\alpha/2} = 2.576$

$1-\alpha = 0.95,\ z_{\alpha/2} = 1.96$

$1-\alpha = 0.90,\ z_{\alpha/2} = 1.645$

$\frac{1}{n}$ $\frac{1}{n-1}$

$X_i \to 0, 1$

$X_i^2 \to 0, 1$

Confidence interval for μ - Bootstrap estimate

When the true σ^2 is not known, plug in the sample variance.

$n = 30$

$n \to \infty$

Corollary D. Under the i.i.d assumption,

$$P\left(\bar{X}_n - \frac{z_{\alpha/2}\hat{\sigma}_n}{\sqrt{n}} \le \mu \le \bar{X}_n + \frac{z_{\alpha/2}\hat{\sigma}_n}{\sqrt{n}}\right) \approx 1 - \alpha.$$

$$\bar{X}_n \pm \frac{z_{\alpha/2}\hat{\sigma}_n}{\sqrt{n}} \leftarrow \hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}_n^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i - \bar{X}_n^2 = 0.31 - 0.31^2 = 0.214.$$

$$= 0.31 \pm \frac{2.576 \times 0.214}{\sqrt{100}}$$

$$\hat{\sigma}_n = \sqrt{0.214} = 0.462.$$

$$= [0.1909, 0.4290].$$

$$\hat{\sigma}_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

is also appropriate.

$$1 - \alpha = 0.99,\ z_{\alpha/2} = 2.576$$
$$1 - \alpha = 0.95,\ z_{\alpha/2} = 1.96$$
$$1 - \alpha = 0.90,\ z_{\alpha/2} = 1.645$$

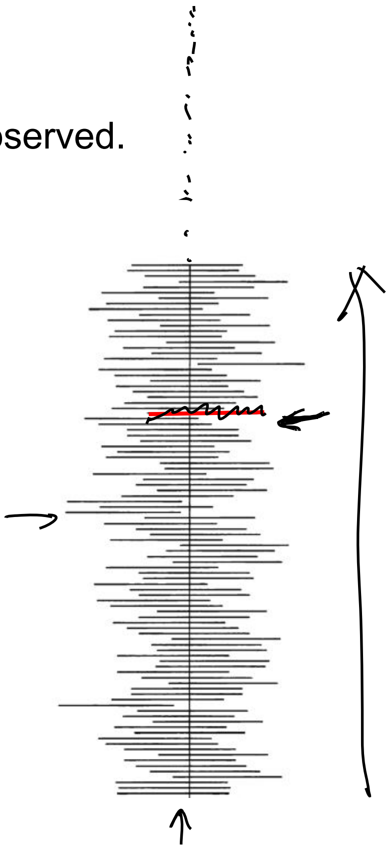
*Compared to the conservative CI: [0.1812, 0.4388].

Confidence interval for μ - Bootstrap estimate

A coin lands heads with probability p. It is tossed 100 times, and 41 heads was observed. The 99% bootstrap confidence interval for p is [0.1909, 0.4290].

Which of the following statement is true?

- There is a 99% probability that p lies within the interval [0.1909, 0.4290].
  False. The confidence level represents theoretical long-run frequency. [0.1909, 0.4290] is calculated from one experiment.
- ~~99%~~ of all possible sample means/proportion fall within [0.1909, 0.4290].
  100 %
  False. The bootstrap CI $\left[\bar{X}_n - \frac{z_{\alpha/2}\sigma}{\sqrt{n}}, \bar{X}_n + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right]$ *always contains the sample mean.*
- There is a 99% probability that a sample mean/proportion from a repeat of the experiment falls within this interval.
  False. Again 99% represents the long-run frequency.
- There is a 99% probability that confidence intervals from all future experiment encompasses the true p.
  True. This is a correct interpretation because it involves numerous repeated sampling.



Point estimation

Chapter 8 of Rice

06/23/2021

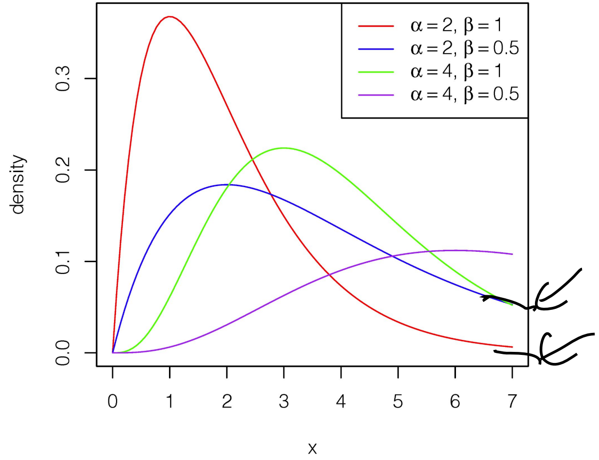
How to estimate parameters other than μ and σ^2?

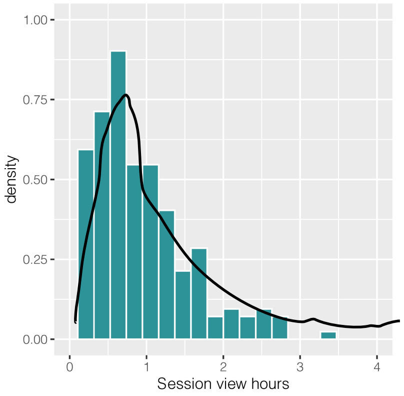
$\bar{X}_n$

$\hat{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$

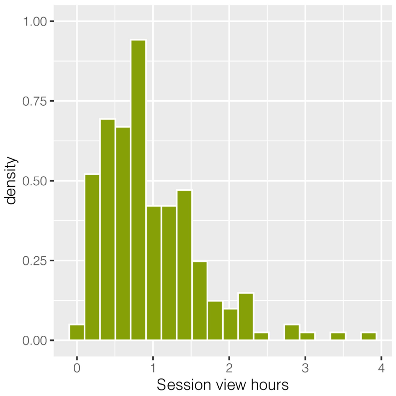
Suppose $X_1, \ldots, X_n$ are i.i.d samples from $\text{Gamma}(\alpha, \beta)$.

In this case , $\mu = \frac{\alpha}{\beta}$, $\sigma^2 = \frac{\alpha}{\beta^2}$.





Design A



Design B

Differences in the distributions between A and B should be reflected in differences in the parameters α and β.

Method of Moments

~~Method of moments~~

$X_1, \cdots X_n$ iid from a population $f(x)$.

Suppose X has the distribution $f(x)$, and $X_1, \ldots, X_n$ are i.i.d samples from that distribution.

Population moments

Population Moments

$EX = \mu \rightarrow$ mean

$\mu = E(X)$

EX^2

$\mu_2 = E(X^2) \quad \mu_2 = \mu^2 \quad \sigma^2 + \mu^2 \rightarrow$ variance

$\mu_3 = E(X^3)$

$\rightarrow EX^3 = \mu_3 \rightarrow$ shape

$\mu_4 = E(X^4)$

$\rightarrow EX^4 = \mu_4 \rightarrow$ kurtosis

tail heaviness

$EX^5 = \mu_5$

$\vdots$

Sample moments

Sample Moments

$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$\hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$

$\hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$

$\hat{\mu}_3 = \frac{1}{n}\sum_{i=1}^{n} X_i^3$

$\hat{\mu}_3 = \frac{1}{n}\sum_{i=1}^{n} X_i^3$

$\hat{\mu}_4 = \frac{1}{n}\sum_{i=1}^{n} X_i^4$

$\hat{\mu}_4 = \frac{1}{n}\sum_{i=1}^{n} X_i^4$

$\vdots$

$\hat{\mu}_5 = \frac{1}{n}\sum_{i=1}^{n} X_i^5$

$\vdots$

Method of Moments (MM): Match the sample moments with their population counterparts.

Method of moments

If unknown parameters $\theta_1, \ldots, \theta_k$ are not exactly moments of the population, they can be the solutions of a system of equations.

Population

$$\begin{cases} \mu = g_1(\theta_1, \ldots, \theta_k) \\ \mu_2 = g_2(\theta_1, \ldots, \theta_k) \\ \vdots \\ \mu_k = g_k(\theta_1, \ldots, \theta_k) \end{cases}$$

Samples

$$\hat{\mu} = g_1\left(\hat{\theta}_1, \ldots, \hat{\theta}_k\right)$$
$$\hat{\mu}_2 = g_2\left(\hat{\theta}_1, \ldots, \hat{\theta}_k\right)$$
$$\vdots$$
$$\hat{\mu}_k = g_k\left(\hat{\theta}_1, \ldots, \hat{\theta}_k\right)$$

Method of Moments (MM): Match the sample moments with their population counterparts.

$\underline{\text{Gamma}(\alpha, \beta)}$

$$\begin{cases} \mu = \dfrac{\alpha}{\beta} \\ \sigma^2 = \dfrac{\alpha}{\beta^2} = \mu_2 - \mu^2 \end{cases}$$

$$\mu_2 = \frac{\alpha}{\beta^2} + \mu^2 = \frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2}$$

$$\Rightarrow \begin{cases} \hat{\mu} = \dfrac{\hat{\alpha}}{\hat{\beta}} \\ \hat{\mu}_2 = \dfrac{\hat{\alpha}}{\hat{\beta}^2} + \dfrac{\hat{\alpha}^2}{\hat{\beta}^2} \end{cases}$$

Method of moments

Example 3. Suppose $X_1, \ldots, X_n$ are i.i.d samples from $\mathrm{Gamma}(\alpha, \beta)$, find Method of Moments for *α* and *β.*

We need to solve the following system of equations:

$$\begin{cases} \hat{\mu} = \frac{1}{n}\sum X_i = \frac{\hat{\alpha}}{\hat{\beta}} \\ \hat{\mu}_2 = \frac{1}{n}\sum X_i^2 = \frac{\hat{\alpha}}{\hat{\beta}^2} + \frac{\hat{\alpha}^2}{\hat{\beta}^2} \end{cases}$$

Plug the first eq into the second:

$$\hat{\mu}_2 = \frac{\hat{\mu}}{\hat{\beta}} + \hat{\mu}^2 \Rightarrow \frac{1}{\hat{\beta}} = \frac{\hat{\mu}_2 - \hat{\mu}^2}{\hat{\mu}}, \quad \hat{\beta} = \frac{\hat{\mu}}{\hat{\mu}_2 - \hat{\mu}^2} = \frac{\bar{X}_n}{\frac{1}{n}\sum X_i^2 - \bar{X}_n^2}$$

$$= \frac{\bar{X}_n}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$$

$$\hat{\alpha} = \hat{\beta}\hat{\mu} = \frac{\hat{X}_n^2}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}.$$

Method of Moments (MM): Match the sample moments with their population counterparts.

k $\quad$ ● ○ ● ○ ● ○ $\quad$ $= m$

$$P(X=m) = \binom{k}{m} p^m (1-p)^{k-m}$$

Method of moments

Example 4. Suppose $X_1, \ldots, X_n$ are i.i.d samples from $X \sim$ Binomial(k, p). Find the Method of Moments estimators for k and p.

We know from STAT 134 that $E(X) = kp$, $\mathrm{Var}(X) = kp(1-p)$.

Solution.
$$\begin{cases} \mu = kp \\ \mu_2 = \mathrm{var}(X) + [E(X)]^2 = kp(1-p) + k^2p^2 \end{cases}$$

Thus, we need to solve:
$$\begin{cases} \hat{\mu} = \frac{1}{n}\sum_{i=1}^n X_i = kp \\ \hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^n X_i^2 = kp(1-p) + k^2p^2 = \hat{\mu}(1-p) + \hat{\mu}^2 \end{cases}$$

Then,
$$\hat{p} = 1 - \frac{\hat{\mu}_2 - \hat{\mu}^2}{\hat{\mu}} = 1 - \frac{\frac{1}{n}\sum X_i^2 - \bar{X}_n^2}{\bar{X}_n} = \frac{\bar{X}_n - \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2}{\bar{X}_n} < 0$$

$$\hat{k} = \frac{\hat{\mu}}{\hat{p}} = \frac{\bar{X}_n \; \hat{\mu}}{1 - \frac{\frac{1}{n}\sum X_i^2 - \bar{X}_n^2}{\bar{X}_n}} = \frac{\bar{X}_n^2}{\bar{X}_n - \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2}$$

Method of Moments (MM): Match the sample moments with their population counterparts.

⇩

MM estimators are not always meaningful.

Tomorrow...

The sampling distribution of MM estimators: §8.3 cont'd