AUA CS108, Statistics, Fall 2020

Lecture 36

Michael Poghosyan

18 nov 2020

Contents

CI Problem Setting

Assume we have a Random Sample from a Parametric Model $\mathcal{F}_\theta$:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

CI Problem Setting

Assume we have a Random Sample from a Parametric Model $\mathcal{F}_\theta$:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

We take $\alpha \in (0,1)$, and call the number $1-\alpha$ the **confidence level**.

CI Problem Setting

Assume we have a Random Sample from a Parametric Model $\mathcal{F}_\theta$:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

We take $\alpha \in (0,1)$, and call the number $1-\alpha$ the **confidence level**.

Problem: Using the Random Sample, construct an interval containing the Unknown Parameter value with Probability not less than $1-\alpha$.

CI Problem Setting

Assume we have a Random Sample from a Parametric Model $\mathcal{F}_\theta$:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

We take $\alpha \in (0, 1)$, and call the number $1 - \alpha$ the **confidence level**.

Problem: Using the Random Sample, construct an interval containing the Unknown Parameter value with Probability not less than $1 - \alpha$.

The usual values of the confidence level are $90\%, 95\%, 99\%$, so the usual values of α are $0.1, 0.05$ and 0.01.

CI

Definition: Assume $0 < \alpha < 1$, and let $L = L(x_1, ..., x_n, \alpha)$, $U = U(x_1, ..., x_n, \alpha)$ be two functions with $L(x_1, ..., x_n, \alpha) \leq U(x_1, ..., x_n, \alpha)$ for all $(x_1, ..., x_n, \alpha)$.

CI

Definition: Assume $0 < \alpha < 1$, and let $L = L(x_1, ..., x_n, \alpha)$, $U = U(x_1, ..., x_n, \alpha)$ be two functions with $L(x_1, ..., x_n, \alpha) \leq U(x_1, ..., x_n, \alpha)$ for all $(x_1, ..., x_n, \alpha)$. The random interval

$$(L, U) = \Big(L(X_1, ..., X_n, \alpha) \ , \quad U(X_1, ..., X_n, \alpha) \Big)$$

is called a **confidence interval (or confidence interval estimator) for θ of confidence level** $1 - \alpha$, if for any $\theta \in \Theta$,

$$\mathbb{P}(L < \theta < U) \geq 1 - \alpha.$$

CI

In the case we have a realization/observation of $X_1, ..., X_n$, say, $x_1, ..., x_n$, then the interval

$$\Big(L(x_1, ..., x_n, \alpha),\ \ U(x_1, ..., x_n, \alpha)\Big)$$

will be an **interval estimate for θ for the confidence level** $(1 - \alpha)$.

CI

In the case we have a realization/observation of $X_1, ..., X_n$, say, $x_1, ..., x_n$, then the interval

$$\Big(L(x_1, ..., x_n, \alpha),\ \ U(x_1, ..., x_n, \alpha)\Big)$$

will be an **interval estimate for θ for the confidence level** $(1 - \alpha)$.

Going back to our CI, CI of the confidence level $1 - \alpha$ is a Random Interval that contains θ in more than $(1 - \alpha) \cdot 100\%$ of cases.

CI, Interpretation

Note: It is important to understand, that in the CI definition

$$\mathbb{P}(L < \theta < U) \geq 1 - \alpha$$

θ is not our r.v., θ is our unknown constant Parameter, so we do not read this as "with high Probability, θ is in (L, U)".

CI, Interpretation

Note: It is important to understand, that in the CI definition

$$\mathbb{P}(L < \theta < U) \geq 1 - \alpha$$

θ is not our r.v., θ is our unknown constant Parameter, so we do not read this as "with high Probability, θ is in (L, U)". Instead, the Interval (L, U) is Random, so we read this as **with high Probability, the interval (L, U) will contain θ.**

---

[1]But not different θ-s !!

CI, Interpretation

Note: It is important to understand, that in the CI definition

$$\mathbb{P}(L < \theta < U) \geq 1 - \alpha$$

θ is not our r.v., θ is our unknown constant Parameter, so we do not read this as "with high Probability, θ is in (L, U)". Instead, the Interval (L, U) is Random, so we read this as **with high Probability, the interval (L, U) will contain θ.**

So, if we will have/generate different observations, we will have different Intervals[1] (L, U), and we want to have that most of the time that interval contains our unknown Parameter value.

[1] But not different θ-s !!

CI, **R** Simulation

Example: Consider an example: our Model is $Exp(\lambda)$, and we have an observation from it. Let us take a Random Sample for the general case: $X_1, X_2, ..., X_n$ from $Exp(\lambda)$.

CI, **R** Simulation

Example: Consider an example: our Model is $Exp(\lambda)$, and we have an observation from it. Let us take a Random Sample for the general case: $X_1, X_2, ..., X_n$ from $Exp(\lambda)$. Then good **Estimator** for λ is

CI, **R** Simulation

Example: Consider an example: our Model is $Exp(\lambda)$, and we have an observation from it. Let us take a Random Sample for the general case: $X_1, X_2, ..., X_n$ from $Exp(\lambda)$. Then good **Estimator** for λ is

$$\hat{\lambda} = \frac{1}{\overline{X}}.$$

CI, **R** Simulation

Example: Consider an example: our Model is $Exp(\lambda)$, and we have an observation from it. Let us take a Random Sample for the general case: $X_1, X_2, ..., X_n$ from $Exp(\lambda)$. Then good **Estimator** for λ is

$$\hat{\lambda} = \frac{1}{\overline{X}}.$$

Now, let us take as CI

CI, **R** Simulation

Example: Consider an example: our Model is $Exp(\lambda)$, and we have an observation from it. Let us take a Random Sample for the general case: $X_1, X_2, ..., X_n$ from $Exp(\lambda)$. Then good **Estimator** for λ is

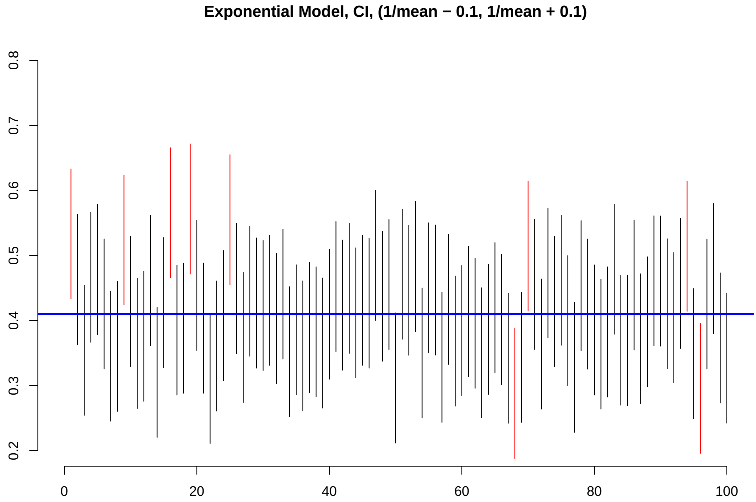
$$\hat{\lambda} = \frac{1}{\overline{\overline{X}}}.$$

Now, let us take as CI

$$\left(\frac{1}{\overline{\overline{X}}} - 0.1, \frac{1}{\overline{\overline{X}}} + 0.1\right)$$

and do some simulations:

CI, **R** Simulation



CI, **R** Simulation, Code

```r
#CI Idea, Exponential Model
lambda <- 0.41
conf.level <- 0.95; a = 1 - conf.level
sample.size <- 50; no.of.intervals <- 100
epsilon <- 0.1

plot.new()
plot.window(xlim = c(0,no.of.intervals), ylim = c(0.2,0.8))
axis(1); axis(2)
title("Exponential Model, CI, (1/mean - 0.1, 1/mean + 0.1)")
for(i in 1:no.of.intervals){
  x <- rexp(sample.size, rate = lambda)
  lo <- 1/mean(x) - epsilon; up <- 1/mean(x) + epsilon
  if(lo > lambda || up < lambda){
    segments(c(i), c(lo), c(i), c(up), col = "red")
  }
  else{
    segments(c(i), c(lo), c(i), c(up))
  }
}
abline(h = lambda, lwd = 2, col = "blue")
```

Methods to obtain Confidence Intervals

We will consider several methods to construct CIs:

- Chebyshev Inequality Based;
- Pivotal Quantity Based

CI for the Mean, Variance is given, Cheby Method

Example: Assume $X_1, X_2, ..., X_n$ are Independent r.v. with the same Mean $\mathbb{E}(X_k) = \mu$ and the same Variance $Var(X_k) = \sigma^2$.

CI for the Mean, Variance is given, Cheby Method

Example: Assume $X_1, X_2, ..., X_n$ are Independent r.v. with the same Mean $\mathbb{E}(X_k) = \mu$ and the same Variance $Var(X_k) = \sigma^2$. Assume σ^2 is known, and μ is unknown, and we want to estimate it using the Sample.

CI for the Mean, Variance is given, Cheby Method

Example: Assume $X_1, X_2, ..., X_n$ are Independent r.v. with the same Mean $\mathbb{E}(X_k) = \mu$ and the same Variance $Var(X_k) = \sigma^2$. Assume σ^2 is known, and μ is unknown, and we want to estimate it using the Sample. The problem is to construct a CI for μ of given Confidence Level $1 - \alpha$.

CI for the Mean, Variance is given, Cheby Method

Example: Assume $X_1, X_2, ..., X_n$ are Independent r.v. with the same Mean $\mathbb{E}(X_k) = \mu$ and the same Variance $Var(X_k) = \sigma^2$. Assume σ^2 is known, and μ is unknown, and we want to estimate it using the Sample. The problem is to construct a CI for μ of given Confidence Level $1 - \alpha$.

By the Chebyshev inequality method, we can obtain that the interval

$$\left(\overline{X} - \frac{\sigma}{\sqrt{n \cdot \alpha}},\ \overline{X} + \frac{\sigma}{\sqrt{n \cdot \alpha}}\right)$$

is a CI for μ of Confidence Level $1 - \alpha$.

CI for the Mean, Variance is given, Cheby Method

Example: Assume $X_1, X_2, ..., X_n$ are Independent r.v. with the same Mean $\mathbb{E}(X_k) = \mu$ and the same Variance $Var(X_k) = \sigma^2$. Assume σ^2 is known, and μ is unknown, and we want to estimate it using the Sample. The problem is to construct a CI for μ of given Confidence Level $1 - \alpha$.

By the Chebyshev inequality method, we can obtain that the interval

$$\left(\overline{X} - \frac{\sigma}{\sqrt{n \cdot \alpha}},\ \overline{X} + \frac{\sigma}{\sqrt{n \cdot \alpha}}\right)$$

is a CI for μ of Confidence Level $1 - \alpha$.

Note: Here

$$\frac{\sigma}{\sqrt{n \cdot \alpha}}$$

is called the **Margin of Error** (for the Interval Estimate of μ, given σ^2).

Some Notes

Two notes about the obtained CI - in fact, these notes will work also for other cases too:

Some Notes

Two notes about the obtained CI - in fact, these notes will work also for other cases too:

The CI length obtained above is

$$\frac{2\sigma}{\sqrt{n \cdot \alpha}}.$$

Note: If we increase n, the CI gets narrower. This is intuitive: if we collect more data, we can estimate the parameter more precisely, we can enclose it in a smaller length interval.

Some Notes

Two notes about the obtained CI - in fact, these notes will work also for other cases too:

The CI length obtained above is

$$\frac{2\sigma}{\sqrt{n \cdot \alpha}}.$$

Note: If we increase n, the CI gets narrower. This is intuitive: if we collect more data, we can estimate the parameter more precisely, we can enclose it in a smaller length interval.

Note: If we increase the Confidence Level, i.e., if we decrease α, then the length of CI increases. This is intuitive too: if we want to be more sure where our unknown Parameter is lying, we will get a larger interval.

CI for the Proportion, Cheby Method

Example: Now, let us construct a CI of CLevel $1 - \alpha$ for p in the *Bernoulli*(p) Model.

CI for the Proportion, Cheby Method

Example: Now, let us construct a CI of CLevel $1-\alpha$ for p in the *Bernoulli*(p) Model.We assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim Bernoulli(p).$$

CI for the Proportion, Cheby Method

Example: Now, let us construct a CI of CLevel $1-\alpha$ for p in the *Bernoulli*(p) Model.We assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathit{Bernoulli}(p).$$

The CI for p by Chebyshev Inequality will be

$$\left(\overline{X} - \frac{1}{2\sqrt{n\cdot\alpha}},\ \overline{X} + \frac{1}{2\sqrt{n\cdot\alpha}}\right)$$

is a CI for p of level $1-\alpha$.

CI for the Proportion, Cheby Method

Example: Now, let us construct a CI of CLevel $1-\alpha$ for p in the *Bernoulli*(p) Model.We assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathit{Bernoulli}(p).$$

The CI for p by Chebyshev Inequality will be

$$\left(\overline{X} - \frac{1}{2\sqrt{n\cdot\alpha}},\ \overline{X} + \frac{1}{2\sqrt{n\cdot\alpha}}\right)$$

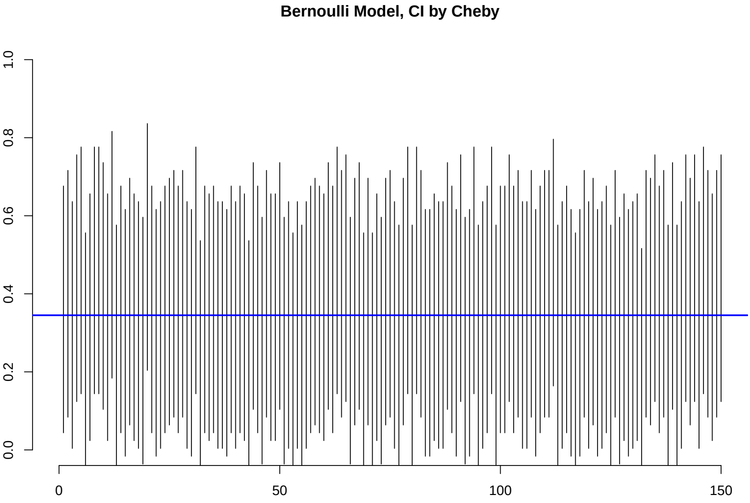
is a CI for p of level $1-\alpha$.

Note: Here

$$\frac{1}{2\sqrt{n\cdot\alpha}}$$

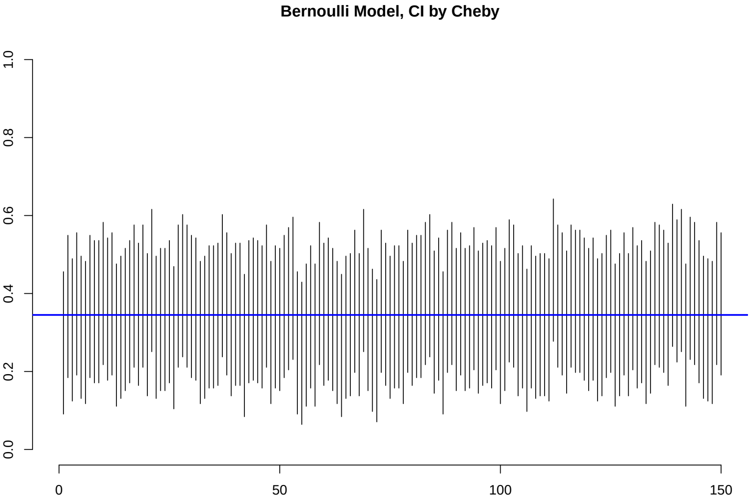
is called the **Margin of Error** (for the Interval Estimate of p).

CI for Bernoulli, **R** Simulation



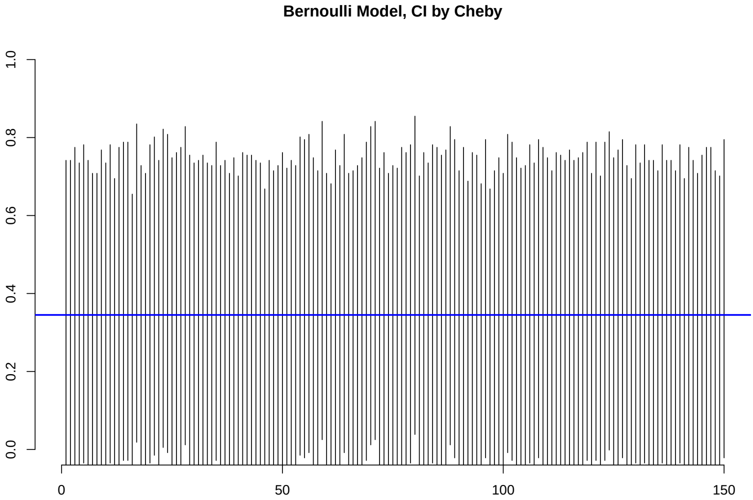
Sample Size $= 50,\ CL = 95\%$

CI for Bernoulli, **R** Simulation



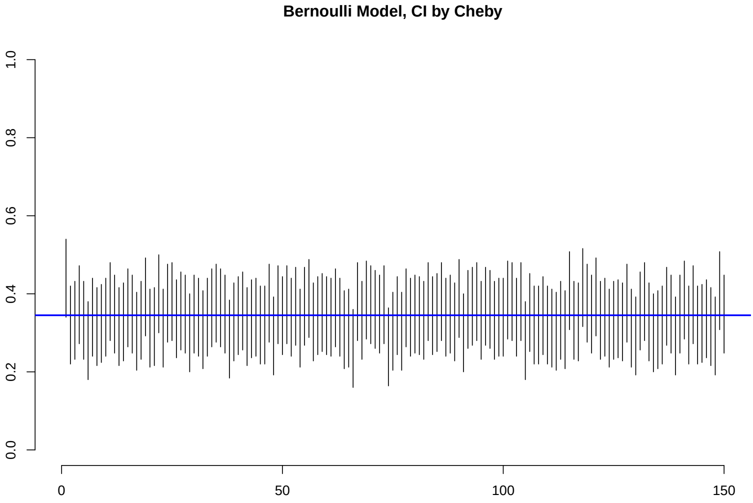
Sample Size $= 150$, $CL = 95\%$

CI for Bernoulli, **R** Simulation



Sample Size $= 150$, $CL = 99\%$

CI for Bernoulli, **R** Simulation



Sample Size $= 250$, $CL = 90\%$

CI, **R** Simulation, Code

```r
#CI Idea, Bernoulli Model
p <- 0.345
conf.level <- 0.9; a = 1 - conf.level
sample.size <- 250; no.of.intervals <- 150
ME <- 1/(2*sqrt(sample.size*a)) #Margin of Error
plot.new()
plot.window(xlim = c(0,no.of.intervals), ylim = c(0,1))
axis(1); axis(2)
title("Bernoulli Model, CI by Cheby")
for(i in 1:no.of.intervals){
  x <- rbinom(sample.size, size = 1, prob = p)
  lo <- mean(x) - ME
  up <- mean(x) + ME
  if(lo > p || up < p){
    segments(c(i), c(lo), c(i), c(up), col = "red")
  }
  else{
    segments(c(i), c(lo), c(i), c(up))
  }
}
abline(h = p, lwd = 2, col = "blue")
```