CS70: Jean Walrand: Lecture 29.

Confidence Intervals

1. Confidence?
2. Example
3. Review of Chebyshev
4. Confidence Interval with Chebyshev
5. More examples

Confidence?

- You flip a coin once and get H.

  Do think that $Pr[H] = 1$?
- You flip a coin 10 times and get 5 Hs.

  Are you sure that $Pr[H] = 0.5$?
- You flip a coin 10^6 times and get 35% of Hs.

  How much are you willing to bet that $Pr[H]$ is exactly 0.35?

  How much are you willing to bet that $Pr[H] \in [0.3, 0.4]$?

More generally, you estimate an unknown quantity θ.

Your estimate is $\hat{\theta}$.

How much confidence do you have in your estimate?

Confidence?

Confidence is essential is many applications:

- How effective is a medication?
- Are we sure of the milage of a car?
- Can we guarantee the lifespan of a device?
- We simulated a system. Do we trust the simulation results?
- Is an algorithm guaranteed to be fast?
- Do we know that a program has no bug?

As scientists and engineers, you should become convinced of this fact:

An estimate without confidence level is useless!

Confidence Interval

The following definition captures precisely the notion of confidence.

Definition: Confidence Interval

An interval $[a, b]$ is a 95%-confidence interval for an unknown quantity θ if

$$Pr[\theta \in [a, b]] \geq 95\%.$$

The interval $[a, b]$ is calculated on the basis of observations.

Here is a typical framework. Assume that $X_1, X_2, \ldots, X_n$ are i.i.d. and have a distribution that depends on some parameter θ.

For instance, $X_n = B(\theta)$.

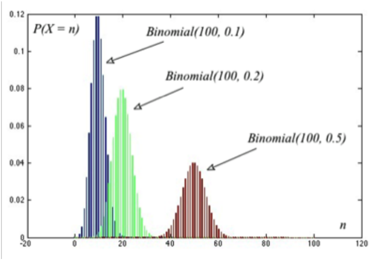
Thus, more precisely, given θ, the random variables X_n are i.i.d. with a known distribution (that depends on θ).

- We observe $X_1, \ldots, X_n$
- We calculate $a = a(X_1, \ldots, X_n)$ and $b = b(X_1, \ldots, X_n)$
- If we can guarantee that $Pr[\theta \in [a, b]] \geq 95\%$, then $[a, b]$ is a 95%-CI for θ.

Confidence Interval: Applications

- We poll 1000 people.
  - Among those, 48% declare they will vote for Trump.
  - We do some calculations ....
  - We conclude that $[0.43, 0.53]$ is a 95%-CI for the fraction of all the voters who will vote for Trump. (Arghhh.)
- We observe $1,000$ heart valve replacements that were performed by Dr. Bill.
  - Among those, 35 patients died during surgery. (Sad example!)
  - We do some calculations ...
  - We conclude that $[1\%, 5\%]$ is a 95%-CI for the probability of dying during that surgery by Dr. Bill.
  - We do a similar calculation for Dr. Fred.
  - We find that $[8\%, 12\%]$ is a 95%-CI for Dr. Fred's surgery.
  - What surgeon do you choose?

Coin Flips: Intuition



Say that you flip a coin $n = 100$ times and observe 20 Hs.

If $p := Pr[H] = 0.5$, this event is very unlikely.

Intuitively, if is unlikely that the fraction of Hs, say A_n, differs a lot from $p := Pr[H]$.

Thus, it is unlikely that p differs a lot from A_n. Hence, one should be able to build a confidence interval $[A_n - \delta, A_n + \delta]$ for p.

The key idea is that $|A_n - p| \leq \delta \Leftrightarrow p \in [A_n - \delta, A_n + \delta]$.

Thus, $Pr[|A_n - p| > \delta] \leq 5\% \Leftrightarrow Pr[p \in [A_n - \delta, A_n + \delta]] \geq 95\%$.

It remains to find δ such that $Pr[|A_n - p| > \delta] \leq 5\%$.

One approach: Chebyshev.

Confidence Interval with Chebyshev

- Flip a coin n times. Let A_n be the fraction of Hs.
- Can we find δ such that $Pr[|A_n - p| > \delta] \leq 5\%$?

Using Chebyshev, we will see that $\delta = 2.25\frac{1}{\sqrt{n}}$ works. Thus

$$[A_n - \frac{2.25}{\sqrt{n}}, A_n + \frac{2.25}{\sqrt{n}}] \text{ is a 95\%-CI for } p.$$

Example: If $n = 1500$, then $Pr[p \in [A_n - 0.05, A_n + 0.05]] \geq 95\%$.

In fact, we will see later that $a = \frac{1}{\sqrt{n}}$ works, so that with $n = 1,500$ one has $Pr[p \in [A_n - 0.02, A_n + 0.02]] \geq 95\%$.

Confidence Intervals: Result

Theorem:
Let X_n be i.i.d. with mean μ and variance σ^2.
Define $A_n = \frac{X_1 + \cdots + X_n}{n}$. Then,

$$Pr[\mu \in [A_n - 4.5\frac{\sigma}{\sqrt{n}}, A_n + 4.5\frac{\sigma}{\sqrt{n}}]] \geq 95\%.$$

Thus, $[A_n - 4.5\frac{\sigma}{\sqrt{n}}, A_n + 4.5\frac{\sigma}{\sqrt{n}}]]$ is a 95%-CI for μ.

Example: Let $X_n = 1\{$ coin n yields $H\}$. Then

$$\mu = E[X_n] = p := Pr[H]. \text{ Also, } \sigma^2 = var(X_n) = p(1-p) \leq \frac{1}{4}.$$

Hence, $[A_n - 4.5\frac{1/2}{\sqrt{n}}, A_n + 4.5\frac{1/2}{\sqrt{n}}]]$ is a 95%-CI for p.

Confidence Interval: Analysis

We prove the theorem, i.e., that $A_n \pm 4.5\sigma/\sqrt{n}$ is a 95%-CI for μ.

From Chebyshev:

$$Pr[|A_n - \mu| \geq 4.5\sigma/\sqrt{n}] \leq \frac{var(A_n)}{[4.5\sigma/\sqrt{n}]^2} = \frac{n}{20\sigma^2} var(A_n).$$

Now,

$$\begin{aligned} var(A_n) &= var(\frac{X_1 + \cdots + X_n}{n}) = \frac{1}{n^2} var(X_1 + \cdots + X_n) \\ &= \frac{1}{n^2} \times n.var(X_1) = \frac{1}{n}\sigma^2. \end{aligned}$$

Hence,

$$Pr[|A_n - \mu| \geq 4.5\sigma/\sqrt{n}] \leq \frac{n}{20\sigma^2} \times \frac{1}{n}\sigma^2 = 5\%.$$

Thus,

$$Pr[|A_n - \mu| \leq 4.5\sigma/\sqrt{n}] \geq 95\%.$$

Hence,

$$Pr[\mu \in [A_n - 4.5\sigma/\sqrt{n}, A_n + 4.5\sigma/\sqrt{n}]] \geq 95\%.$$

$\square$

Confidence interval for p in $B(p)$

Let X_n be i.i.d. $B(p)$. Define $A_n = (X_1 + \cdots + X_n)/n$.

Theorem:

$$[A_n - \frac{2.25}{\sqrt{n}}, A_n + \frac{2.25}{\sqrt{n}}] \text{ is a 95\%-CI for } p.$$

Proof:

We have just seen that

$$Pr[\mu \in [A_n - 4.5\sigma/\sqrt{n}, A_n + 4.5\sigma/\sqrt{n}]] \geq 95\%.$$

Here, $\mu = p$ and $\sigma^2 = p(1-p)$. Thus, $\sigma^2 \leq \frac{1}{4}$ and $\sigma \leq \frac{1}{2}$.

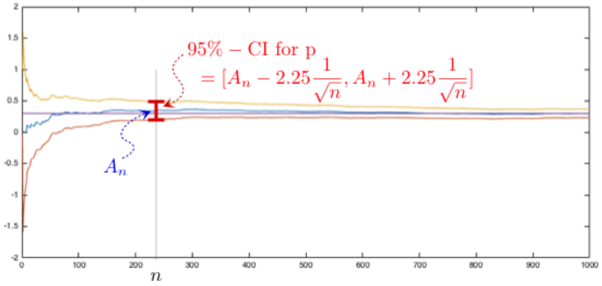
Thus,

$$Pr[\mu \in [A_n - 4.5 \times 0.5/\sqrt{n}, A_n + 4.5 \times 0.5/\sqrt{n}]] \geq 95\%.$$

□

Confidence interval for p in $B(p)$

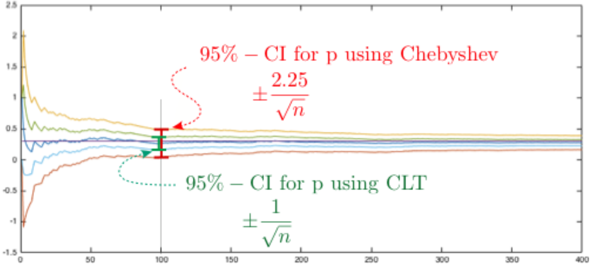
An illustration:



Good practice: You run your simulation, or experiment. You get an estimate. You indicate your confidence interval.

Confidence interval for p in $B(p)$

Improved CI: Later we will see that we can replace 2.25 by 1.



Quite a bit of work to get there: continuous random variables; Gaussian; Central Limit Theorem.

Confidence Interval for $1/p$ in $G(p)$

Let X_n be i.i.d. $G(p)$. Define $A_n = (X_1 + \cdots + X_n)/n$.

Theorem:

$$[\frac{A_n}{1+4.5/\sqrt{n}}, \frac{A_n}{1-4.5/\sqrt{n}}] \text{ is a 95\%-CI for } \frac{1}{p}.$$

Proof: We know that

$$Pr[\mu \in [A_n - 4.5\sigma/\sqrt{n}, A_n + 4.5\sigma/\sqrt{n}]] \geq 95\%.$$

Here, $\mu = \frac{1}{p}$ and $\sigma = \frac{\sqrt{1-p}}{p} \leq \frac{1}{p}$. Hence,

$$Pr[\frac{1}{p} \in [A_n - 4.5\frac{1}{p\sqrt{n}}, A_n + 4.5\frac{1}{p\sqrt{n}}]] \geq 95\%.$$

Now, $A_n - 4.5\frac{1}{p\sqrt{n}} \leq \frac{1}{p} \leq \frac{1}{p} \leq A_n + 4.5\frac{1}{p\sqrt{n}}$ is equivalent to

$$\frac{A_n}{1+4.5/\sqrt{n}} \leq \frac{1}{p} \leq \frac{A_n}{1-4.5/\sqrt{n}}.$$

□

Examples: $[0.7A_{100}, 1.8A_{100}]$ and $[0.96A_{10000}, 1.05A_{10000}]$.

Which Coin is Better?

You are given coin A and coin B. You want to find out which one has a larger $Pr[H]$. Let p_A and p_B be the values of $Pr[H]$ for the two coins.

Approach:

- Flip each coin n times.
- Let A_n be the fraction of Hs for coin A and B_n for coin B.
- Assume $A_n > B_n$. It is tempting to think that $p_A > p_B$. Confidence?

Analysis: Note that

$$E[A_n - B_n] = p_A - p_B \text{ and } var(A_n - B_n) = \frac{1}{n}(p_A(1-p_A) + p_B(1-p_B)) \leq \frac{1}{2n}.$$

Thus, $Pr[|A_n - B_n - (p_A - p_B)| > \delta] \leq \frac{1}{2n\delta^2}$, so

$$Pr[p_A - p_B \in [A_n - B_n - \delta, A_n - B_n + \delta]] \geq 1 - \frac{1}{2n\delta^2}, \text{ and}$$

$$Pr[p_A - p_B \geq 0] \geq 1 - \frac{1}{2n(A_n - B_n)^2}.$$

Example: With $n = 100$ and $A_n - B_n = 0.2$, $Pr[p_A > p_B] \geq 1 - \frac{1}{8} = 0.875$.

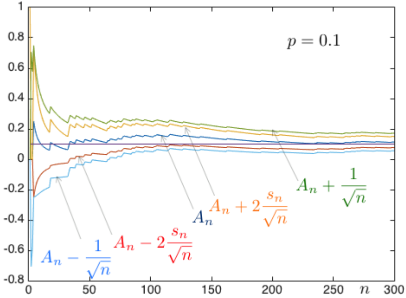
Unknown σ

For $B(p)$, we wanted to estimate p. The CI requires $\sigma = \sqrt{p(1-p)}$. We replaced σ by an upper bound: $1/2$.

In some applications, it may be OK to replace σ^2 by the following sample variance:

$$s_n^2 := \frac{1}{n} \sum_{m=1}^{n} (X_m - A_n)^2.$$

However, in some cases, this is dangerous! The theory says it is OK if the distribution of X_n is nice (Gaussian). This is used regularly in practice. However, be aware of the risk.



Summary

Confidence Intervals

1. Estimates without confidence level are useless!
2. $[a, b]$ is a 95%-CI for θ if $Pr[\theta \in [a, b]] \geq 95\%$.
3. Using Chebyshev: $[A_n - 4.5\sigma/\sqrt{n}, A_n + 4.5\sigma/\sqrt{n}]$ is a 95%-CI for μ.
4. Using CLT, we will replace 4.5 by 2.
5. When σ is not known, one can replace it by an upper bound.
6. Examples: $B(p), G(p)$, which coin is better?
7. In some cases, one can replace σ by the empirical standard deviation.