Course › Unit 2 Foundation of Inference › Lecture 4: Parametric Estimation and Confidence Intervals › 10. Solving for a Confidence Interval

10. Solving for a Confidence Interval

Confidence Interval by Solving for p

We did not make any approximation.

All we did was rewrite what we had in the first place

and solved it for p.

We did not make any approximation.

We did not make anything conservative.

This is as precise as it gets, so long

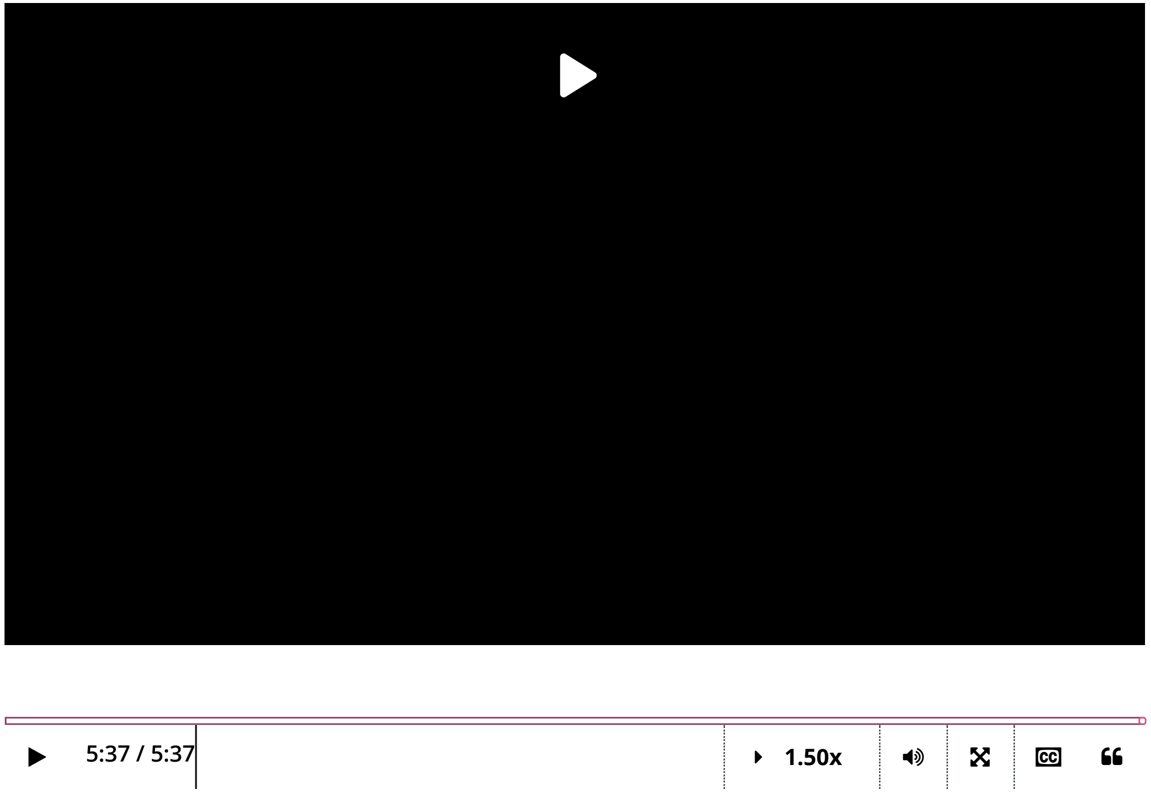
as you believe that the central limit theorem is correct.

But this will certainly give you some things

such that the limit that this confidence interval contains

p is equal to 1 minus alpha, equality being the--

OK?



So it's better than the conservative bound, in a way.

So we'll come back to it.

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Solving for a Confidence Interval: Algebra

2/2 points (graded)

In the problems on this page, we will continue building the confidence interval of asymptotical level **95%** by solving for p as in the video.

Recall that $R_1, \ldots, R_n \overset{iid}{\sim}$ Ber (p) for some unknown parameter p, and we estimate p using the estimator $\hat{p} = \overline{R}_n = \frac{1}{n}\sum_{i=1}^{n} R_i$.

As in the method using a conservative bound, our starting point is the result of the central limit theorem:

$$\lim_{n\to\infty} \mathbf{P}\left(\left|\sqrt{n}\frac{\overline{R}_n - p}{\sqrt{p\,(1-p)}}\right| < q_{\alpha/2}\right) = 1-\alpha.$$

In this second method, we solve for values of p that satisfy the inequality $\left|\sqrt{n}\frac{\overline{R}_n - p}{\sqrt{p\,(1-p)}}\right| < q_{\alpha/2}$.

To do this, we manipulate $\left|\sqrt{n}\frac{\overline{R}_n - p}{\sqrt{p\,(1-p)}}\right| < q_{\alpha/2}$ into an inequality involving a quadratic function $Ap^2 + Bp + C$ where $A > 0,\ B,\ C$ depend on $n, q_{\alpha/2},$ and $\overline{R}_n$. Which of the following is the correct inequality?
(We will use find the values of $A,\ B,$ and C in the next problem.)

● $Ap^2 + Bp + C < 0$ where $A > 0$.

○ $Ap^2 + Bp + C > 0$ where $A > 0$.

✓

Let p_1 and p_2 with $0 < p_1 < p_2 < 1$ be the two roots of the quadratic function $Ap^2 + Bp + C$. What values of p satisfy the correct inequality above?

○ $(p < p_1) \cup (p > p_2)$

● $p_1 < p < p_2$

○ $0 < p < p_1$

○ $p_2 < p < 1$

○ $0 < p < 1$

✓

Solution:

$$\begin{aligned}\left|\sqrt{n}\frac{\overline{R}_n - p}{\sqrt{p\left(1-p\right)}}\right| < q_{\alpha/2} &\implies \left(\sqrt{n}\frac{\overline{R}_n - p}{\sqrt{p\left(1-p\right)}}\right)^2 < q_{\alpha/2}^2 \\ &\implies \left(\overline{R}_n - p\right)^2 < \frac{p\left(1-p\right)q_{\alpha/2}^2}{n} \\ &\implies p^2\left(1+\frac{q_{\alpha/2}^2}{n}\right) - p\left(2\overline{R}_n + \frac{q_{\alpha/2}^2}{n}\right) + \left(\overline{R}_n\right)^2 < 0\end{aligned}$$

Hence, the inequality is of the form $Ap^2 + Bp + C{<}0$ for some $A > 0$.
The quadratic function $Ap^2 + Bp + C < 0$ where $A > 0$ is convex, so the parabola opens up, and the region in which the parabola is below the x-axis is the interval between the two roots. Given $0 < p_1 < p_2 < 1$, the region is $p_1 < p < p_2$.

Submit You have used 1 of 1 attempt

Answers are displayed within the problem

Solving for a Confidence Interval: Numerical Descriptions

2/2 points (graded)

Continuing from above, enter numerical values for $A > 0$, B, C such that the inequality in the previous problem is equivalent to $\left|\sqrt{n}\frac{\overline{R}_n - p}{\sqrt{p\left(1-p\right)}}\right| < q_{\alpha/2}$ for the case when the sample size is $n = 100$, and the observed value of $\overline{R}_n$ is 0.645.

Carry out the computations with the goal of computing a confidence interval of p at asymptotic level 95%. **Note:** Because polynomials differing by only an overall rescaling constant yield the same roots, use $C = \left(\overline{R}_n\right)^2$ here as in the previous problem.

(If necessary, round your answers to the nearest four decimal places (10^{-4}).

$0 < A =$ 1.0384145882069413 ✓ **Answer:** 1+(1.96^2)/100

$B =$ -1.3284145882069414 ✓ **Answer:** -(2*0.645+1.96^2/100)

$C =$ 0.41602500000000003 ✓ **Answer:** 0.645^2

Now, as indicated previously, use the above values (**rounded to the nearest** 10^{-4}) to compute a confidence interval $\mathcal{I}_{\text{solve}}$ of p of asymptotic level 95%.
If necessary, round your endpoints to the nearest two decimal places (10^{-2}).

$p \in [$ 0.54744397 ✓ **Answer:** 0.5473323 , 0.73182791 ✓ **Answer:** 0.7319435 $]$

Solution:

Recall from the previous problem that

$$\left|\sqrt{n}\frac{\overline{R}_n - p}{\sqrt{p\,(1-p)}}\right| < q_{\alpha/2} \implies p^2\left(1 + \frac{q_{\alpha/2}^2}{n}\right) - p\left(2\overline{R}_n + \frac{q_{\alpha/2}^2}{n}\right) + \left(\overline{R}_n\right)^2 < 0.$$

Plugging $n = 100$, $\overline{R}_n = 0.645$, and $q_{\alpha/2} = q_{0.025} = 1.96$ into the inequality above gives

$$p^2\left(1 + \frac{1.96^2}{100}\right) - p\left(2\,(0.645) + \frac{1.96^2}{100}\right) + 0.645^2 < 0.$$

The quadratic formula gives the roots $\frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$, which are

$$\begin{aligned} p_1 &= 0.5473323 \\ p_2 &= 0.7319435. \end{aligned}$$

This gives the confidence interval $[p_1, p_2] \approx [0.55, 0.73]$.

Submit

You have used 1 of 3 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 2 Foundation of Inference:Lecture 4: Parametric Estimation and Confidence Intervals / 10. Solving for a Confidence Interval

Add a Post

All Posts

quartile computation

discussion posted a day ago by anonymous

My calculation is mostly wrong. What is the quartile or q_alpha/2. I can look this up, from a table, I'm not sure whether this is the Z related to the alpha of 0.05 or that it should be 0.025?

This post is visible to everyone.

Add a Response

2 responses

Erocha

a day ago

Slides 19 and 21 should help:

https://courses.edx.org/asset-v1:MITx+18.6501x+3T2018+type@asset+block@lectureslides_Chap2annotlast-1.pdf

Add a comment

davidstuartbruce

about 2 hours ago

In general terms, a 95% CI is often considered significant and is roughly equivalent to "2 standard deviations from the mean". But as Prof. Rigollet humorously mentioned, after this course we can "impress" people at parties by pedantically pointing out that the relevant value is actually 1.96, and for even more statistical geek cred one can memorize the quantiles for other levels of significance.

If you have R, qnorm() is the relevant function for finding $q_{\alpha/2}$, though be careful with the sign convention. In python, the corresponding function appears to be scipy.stats.norm.ppf, though I haven't used python much for statistics.

Regards, David

Add a comment

Showing all responses

Add a response:

Preview

Submit

Learn About Verified Certificates