# asymptotic and exact confidence interval



0

Assume we toss a thumbtack 300 times. After every time, we note 1 if it lands point up or 0 if it lands point down. In summary, we get 124 times 1.

So we know that the number of the rounds with outcome 1 is $\sim Bin(300, p)$ with unknown parameter $p$. Furthermore, $\mathcal{X} := \{0,1\}^{300}$, $\Theta := [0,1]$, $p \in \Theta$ (is this correct?)

Now there is the task:

Define the term 'asymptotic and exact confidence interval' for the niveau $1-\alpha > 0$. Give a 95% confidence interval for the probability that the thumb will land point up.

I do have the formal definitions of the asymptotic and exact confidence interval, but I don't really understand it. Could anyone explain it to me referring to this specific example?

The definitions are:

$Definition$

Let $(\mathbb{P}_\theta)_{\theta\in\Theta}$ be a statistical model with $\Theta \subset \mathbb{R}^n$ on the sample space $\mathcal{X}$. A reel parameter is a mapping $\gamma : \Theta \to \mathbb{R}$. An interval-valued mapping

$$I : \mathcal{X} \to \mathcal{P}(\mathbb{R}), I(x) = [U(x), O(x)]$$

with the statistics $U, O : \mathcal{X} \to \mathbb{R}$ with $U \leq O$ is called an interval estimation for the parameter $\gamma$

$Definition$

The coverage probability of an interval estimation $I$ for a parameter $\gamma$ is the mapping

$$\theta \to \mathbb{P}_\theta(\{x \in \mathcal{X} : \gamma(\theta) \in I(x)\}), \theta \in \Theta$$

A confidence niveau of an interval estimation is the minimal coverage probability

$$\inf_{\theta\in\Theta} \mathbb{P}_\theta(\gamma(\theta) \in I(x))$$

$Definition$

An inverval estimation $I$ is called (exact) confidence interval for the confidence niveau $1-\alpha$ (for a fixed $\alpha \in [0,1]$), if

$$\forall\theta \in \Theta : \mathbb{P}_\theta(\gamma(\theta) \in I(x)) \geq 1-\alpha$$

$Definition$

$$\forall \theta \in \Theta : \liminf_{n\to\infty} P_\theta^{\otimes n}(\{x \in \mathcal{X}^n : \gamma(\theta) \in I_n(x)\}) \geq 1-\alpha$$

probability statistics hypothesis-testing confidence-interval

edited Feb 18 '18 at 12:24

asked Feb 18 '18 at 9:52
newbie
323 2 14

Can you add the definitions you have to your question? That way it'll help to answer you with the notation / concepts you are familiar with. – owen88 Feb 18 '18 at 10:59

Just done, thanks for the suggestion! – newbie Feb 18 '18 at 12:25

## 1 Answer

1

In your context, you are looking to define a confidence interval for the parameter $p$ associated with a Bernoulli distribution (i.e. the *true* probability $p$ that a thumbtack will land point up).

Fortunately, due to the relationship between Bernoulli and Binomial variables, as you observe this is equivalent to finding a confidence interval for the parameter $p$ of a $\mathrm{Bin}(n,p)$ distribution (with $n=300$ in your instance), based on the observed outcome ($X=124$, in your instance).

For a Binomial distribution, there is one standard example of an exact (1-$\alpha$) confidence interval, called the Clopper-Pearson interval. This has a rather messy formula, and is given by

$$I_\alpha = \left(B\left(\tfrac{\alpha}{2}; X; n-X+1\right),\ B\left(1-\tfrac{\alpha}{2}; X+1; n-X\right)\right),$$

here $B(r\,; v, w)$ denotes the percentile function of a Beta distribution with shape parameters $v, w$. For you, I'd imagine what this function is doesn't matter. In your particular instance the interval is at $\alpha = 0.05$ (i.e. a 95% confidence interval)

$$I_\alpha = \left(B\left(0.975; 125, 177\right), B\left(0.025; 124, 176\right)\right) = (0.3570, 0.4714).$$

This is an *exact* confidence interval: which means that it is guaranteed that at least 95% of the time the true parameter will lie within the interval you calculate.

As an example of a asymptotic confidence interval we can use the standard Normal approximation to the binomial distribution, and the associated confidence interval. Denoting $\hat{p} = X/n$, this interval is given by

$$I = \hat{p} \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{\ }}$$

$$J_\alpha = (0.3576, 0.4691).$$

The difference with this interval is that we cannot say for certain that 95% of the time the result will lie in this interval. In particular when $n$ is small this will not be true, but as $n$ gets large it becomes increasingly close to being true that 95% of observations would fall in the interval. To see why this formula doesn't work for small $n$, suppose that we know $p = 1/2$, and suppose we make one throw, $n = 1$. If this lands point up then the interval we would obtain (from the above formula) would be $J_\alpha = [1, 1]$, whilst if it didn't land point up it would be $J_\alpha = [0, 0]$. In either case, the probability that the true value falls in the interval $J_\alpha$ is clearly 0 (since $p = 1/2$). i.e. the answer does not fall into the interval 95% of the time.

answered Feb 18 '18 at 19:29
owen88
**3,750** 11 21