Central limit theorem for the variance

Asked 2 years, 4 months ago Active 2 years, 4 months ago Viewed 501 times

0

1

The central limit theorem establishes that the average of n i.i.d. random variables tends to a normal law, with parameters μ and σ^2/n.

The average is a non-biaised estimator of the mean of the distribution. If we turn to the case of the non-biased estimator of the distribution variance, s^2, the central limit theorem can also be used as s^2 is the difference of

- the average of n i.i.d. random variables $(\overline{x^2})$,
- the square of a normally distributed variable $(\overline{x}^2)$.

Anyway the computation is uneasy as these two variables aren't independent.

As s^2 is unbiaised, its mean is σ^2. But what is the variance of s^2 ? And is its distribution normal in the limit ?

probability-distributions central-limit-theorem estimation-theory

edited Jun 16 '17 at 20:59

asked Jun 16 '17 at 20:50
Yves Daoust
148k 11 89 250

1 You are considering

$$\bar{x}=\frac{1}{n}\sum_{i=1}^n x_i \qquad s^2=\frac{1}{n}\left(\sum_{i=1}^n x_i^2\right)-\bar{x}^2$$

then

$$E(s^2)=\frac{n-1}{n}\sigma^2$$

hence s^2 is a biased estimator of the variance $\sigma^2=\text{var}(x_i^2)$. – Did Jun 16 '17 at 21:05

@Did: I am unsure about the notation and I want to refer to the unbiaised estimator. Anyway, this distinction is not essential as they are proportional to each other. – Yves Daoust Jun 16 '17 at 21:07

1 Then correct your post. And what prevents you to compute $E(s^4)$? – Did Jun 16 '17 at 21:08

@Did: no idea of the relation between $E(s^4)$, presumably drawn from the moments of the original distribution, and the variance of the limit distribution of s^2 (which I don't know). – Yves Daoust Jun 16 '17 at 21:14

1 Answer

1

Let $s^2 = \frac{n}{n-1}\left(\overline{x^2} - (\overline{x})^2\right) = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x})^2$. Define centered r.v.'s $y_i = x_i - \mathbb{E}x_1$ and rewrite sample variance in terms of this r.v.'s:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^n (y_i - \overline{y})^2 = \frac{n}{n-1}\left(\overline{y^2} - (\overline{y})^2\right) = \overline{y^2} - (\overline{y})^2 + \frac{s^2}{n}.$$

Note that $\sigma^2 = \mathrm{Var}(x_1) = \mathbb{E}[y_1^2]$.

Find the limiting distribution of $\sqrt{n}\left(s^2 - \sigma^2\right)$:

$$\sqrt{n}\left(s^2 - \sigma^2\right) = \sqrt{n}\left(\overline{y^2} - (\overline{y})^2 + \frac{s^2}{n} - \sigma^2\right) = \sqrt{n}\left(\overline{y^2} - \sigma^2\right) - \sqrt{n}(\overline{y})^2 + \sqrt{n}\frac{s^2}{n}. \qquad (1)$$

Next prove that $\sqrt{n}(\overline{y})^2 \xrightarrow{p} 0$ and $\sqrt{n}\,\frac{s^2}{n} = \frac{s^2}{\sqrt{n}} \xrightarrow{p} 0$ as $n \to \infty$. Indeed, by Slutsky's theorem,

$$\sqrt{n}(\overline{y})^2 = \underbrace{\overline{y}}_{\substack{\downarrow p \\ 0}} \cdot \underbrace{\sqrt{n}\,(\overline{y})}_{\substack{\downarrow d \\ N(0,1)}} \xrightarrow{d} 0 \cdot N(0,1) = 0$$

The convergence in distribution to zero implies the convergence in probability.

Next,

$$\frac{s^2}{\sqrt{n}} = s^2 \cdot \frac{1}{\sqrt{n}} \xrightarrow{p} \sigma^2 \cdot 0 = 0.$$

We obtain that the second and third terms in r.h.s. of (1) tends to zero in probability. Consider the first term:

$$\sqrt{n}\left(\overline{y^2} - \sigma^2\right) = \sqrt{n}\left(\overline{y^2} - \mathbb{E}\left[y_1^2\right]\right) \xrightarrow{d} N(0, \mathrm{Var}(y_1^2)) = N(0, \mathbb{E}\left[y_1^4\right] - \sigma^4).$$

By Slutsky's theorem,

$$\sqrt{n}(s^2 - \sigma^2) \xrightarrow{d} N(0, \mathbb{E}\left[y_1^4\right] - \sigma^4) = N(0, \mathbb{E}\left[(x_1 - \mathbb{E}x_1)^4\right] - \sigma^4). \qquad (2)$$

So, you can say that the limiting distribution of s^2 is normal with mean σ^2 and variance answered Jun 18 '17 at 8:59

NCh
7,339 3 8 26