class07

January 27, 2021 10:21 AM

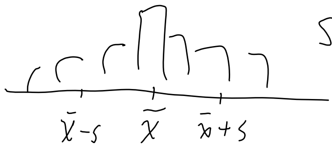
The **standard error** of an estimator $\hat{\theta}$ is its standard deviation $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$. It is the magnitude of a typical or representative deviation between an estimate and the value of θ.

The standard error of an estimator typically involves unknown parameters and thus unknown. Substitution of the estimates of these parameters into $\sigma_{\hat{\theta}}$ yield the **estimated standard error** (of the estimator). The estimated standard error can be denoted either by $\hat{\sigma}_{\hat{\theta}}$ (the $\hat{}$ over σ emphasizes that $\sigma_{\hat{\theta}}$ is being estimated) or by $s_{\hat{\theta}}$.

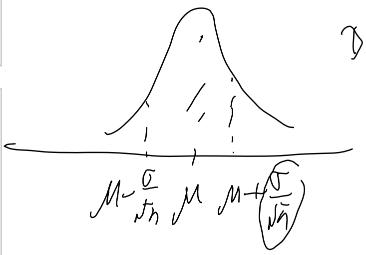
statistic, random variable having a (sampling) dist'n

Pop'n dist'n

Sample dist'n



Dist'n of $\bar{X}$



$\frac{\sigma}{\sqrt{n}}$ ← standard error

Let $\bar{X}$ denote the sample mean of a random sample from a population distribution with mean μ and variance σ^2.

The standard deviation of sample mean $\bar{X}$ is given by $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, which is the standard error (of the sample mean as an estimator of μ).

As σ is unknown, we may estimate it with $s_{\bar{X}} = s/\sqrt{n}$.

$X_1, \cdots, X_{10}$

$5, \cdots, 2$

$\frac{s}{\sqrt{n}}$ ← estimated standard error

$x_1 = 5, x_2 = 4, \bar{x} = 4, s = 2.$

$\hat{\mu} = \frac{X_1+X_2}{2}, \quad \hat{\mu} = \bar{X}, \quad \frac{2}{\sqrt{10}}$

Est'n 4.5, $S = \sqrt{.5}$

$\frac{\sqrt{.5}}{\sqrt{2}}$

E.g., from a population with mean μ, we take a sample of size $n = 25$, $\bar{x} = 1.2$, $s = 0.6$. An estimate of μ is $\hat{\mu} = 1.2$. The (estimated) standard error (of the estimator $\bar{X}$) is $0.6/\sqrt{25} = 0.12$.

(**Note**. Here we have omitted the word "estimated": We estimate μ as 1.2 with a standard error of 0.12.

In the population: (population) standard deviation $= \sigma$ is a parameter; as is a (true) standard error $\sigma_{\hat{\theta}}$.

From data: The (sample) standard deviation s is a fixed number, as an estimate of the population/true mean μ; the (estimated) standard error (of the mean) $s/\sqrt{n}$ is an estimate (much smaller number than s).

Standard error in practice refers to the estimated standard error, e.g. $\frac{S}{\sqrt{n}}$.

Example 7.11 (Example 7.1 continued)
The standard error (of $\hat{p} = Y/n$) is

$$\sigma_{\hat{p}} = \sqrt{V(X/n)} = \sqrt{\frac{V(X)}{n^2}} = \sqrt{\frac{np(1-p)}{n^2}} = \sqrt{\frac{p(1-p)}{n}}$$

A point estimate of p is

$$\hat{p} = x/n = 15/25 = .6$$

The (estimated) standard error (of $\hat{p}$) is

$$\hat{\sigma}_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(.6)(1-.6)}{25}} = .098.$$

$\hat{\sigma}_{\hat{p}}$ $\parallel$ $S_{\hat{p}}$

$X_1, \ldots, X_{25}$.

$X = \begin{cases} 1, & \text{no damage} \\ 0, & \text{otherwise.} \end{cases}$

$E(X) = p = \mu$, $V(X) = \sigma^2 = p(1-p)$,

$\hat{p} = \frac{\sum X_i}{n}$. $E(\hat{p}) = p$.

$V(\hat{p}) = \frac{\sigma^2}{n} = \frac{p(1-p)}{n}$

$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ ← standard error

Estimated SE $= \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Outline

§7.2 Methods of point estimation
The method of moments

If the population is normal with mean μ and standard deviation σ, we may estimate them using their sample counterpart: $\bar{X}$ for μ and S for σ; similarly sample proportion $\hat{p}$ for population proportion p.

How should we estimate the parameters if the population is gamma with shape α and scale β?

Outline

§7.2 Methods of point estimation

The method of moments

The basic idea of this method is to equate certain sample quantities, such as the sample mean, to the corresponding population counterparts. Then solving these equations for unknown parameter values yields the estimators.

$$\sigma^2 = Var(X) = E(X^2) - [E(X)]^2$$
$$= \mu_2 - \mu^2$$

Let $X_1, \ldots, X_n$ be a random sample from a population with pmf or pdf $f(x)$.

Recall: For $k = 1, 2, 3, \ldots$, the k**th moment** (about 0) is

$$\mu_k = E(X^k).$$

We may call it the kth population moment.

The k**th sample moment** is

$$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

$E(X) = \mu = \mu_1$ 1st moment

$E(X^2) = \mu_2 = \mu^2 + \sigma^2 \leftarrow$ 2nd moment

$E(X^3) = \mu_3$, etc.

$M_1 = \frac{1}{n}\sum X_i = \bar{X}$ 1st moment

$M_2 = \frac{1}{n}\sum X_i^2 =$ 2nd moment

Thus the first population moment is $\mu_1 = E(X) = \mu$; and the first sample moment is $M_1 = \bar{X}$.

The second population moment $\mu_2 = E(X^2) = \sigma^2 + \mu^2$; the second sample moment $M_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{n-1}{n}S^2 + \bar{X}^2$.

$$S^2 = \frac{1}{n-1}\sum (X_i - \bar{X})^2$$
$$= \frac{1}{n-1}\left[\sum X_i^2 - n\bar{X}^2\right]$$
$$\sum X_i^2 = (n-1)S^2 + n\bar{X}^2$$
$$\Rightarrow M_2 = \frac{1}{n}\sum X_i^2 = \bar{X}^2 + \frac{n-1}{n}S^2$$

Thus the first population moment is $\mu_1 = E(X) = \mu$; and the first sample moment is $M_1 = \bar{X}$.

The second population moment $\mu_2 = E(X^2) = \sigma^2 + \mu^2$; the second sample moment $M_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{n-1}{n}S^2 + \bar{X}^2$.

$S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})$

$= \frac{1}{n-1}\left[\sum X_i^2 - n\bar{X}^2\right]$

$\sum X_i^2 = (n-1)S^2 + n\bar{X}^2$

$\Rightarrow M_2 = \frac{1}{n}\sum X_i^2 = \bar{X}^2 + \frac{n-1}{n}S^2$

The population moments will be functions of any unknown parameters $\theta_1, \theta_2, \ldots$

> Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with pmf or pdf $f(x; \theta_1, \ldots, \theta_m)$, where $\theta_1, \ldots, \theta_m$ are parameters whose values are unknown. Then the **moment estimators** $\hat{\theta}_1, \ldots, \hat{\theta}_m$ are obtained by equating (the first) m sample moments to the corresponding first m population moments and solving for $\theta_1, \ldots, \theta_m$.

$\mu_1 = E(X) = \frac{1}{n}\sum X_i = M_1$

$\mu_2 = E(X^2) = \frac{1}{n}\sum X_i^2 = M_2$

$\vdots$

Example (Poisson)
Let X be Poisson with parameter k (write "$X \sim$ Poisson(k)" later on). $\theta = k$.

The first moment $E(X) = k$; the first sample moment $M_1 = \bar{X}$.

Equate the sample moment with the theoretical moment:

$$k = M_1 \Rightarrow \hat{k} = M_1 = \bar{X}.$$

When we use the sample mean to estimate the population mean, we are already using the method of moments!

(If the 2nd moment was used, we need to solve $\mu_2 = M_2$, i.e., $k + k^2 = \frac{n-1}{n}S^2 + \bar{X}^2$, resulting another estimator of k.)

For Poisson,
$\mu = k$, $\sigma^2 = k$

$E(X) = k = \frac{1}{n}\sum X_i = \bar{X}$

$\hat{k} = \bar{X}$

$X_1, X_2, \ldots, X_n$

Example (binomial)
Let X be binomial with parameters m and p (write "$X \sim$ binomial(m, p)" later on), where m is the number of trials and p the probability of success.

Need two moments: $E(X) = mp$,
$E(X^2) = \mathsf{Var}(X) + [E(X)]^2 = mp(1-p) + (mp)^2$.

$$\begin{cases} M_1 = \hat{m}\hat{p} \\ M_2 = \hat{m}\hat{p}(1-\hat{p}) + (\hat{m}\hat{p})^2 \end{cases}$$

$$\Rightarrow \hat{p} = 1 - (M_2 - M_1^2)/M_1, \hat{m} = M_1/\hat{p}.$$

(Bernoulli population vs binomial population.)

$E(X) = \underset{\mu}{mp}$, $V(X) = mp(1-p) = \sigma^2$

$E(X^2) = \mu^2 + \sigma^2$

$= (mp)^2 + mp(1-p)$

$\begin{cases} \hat{m}\hat{p} = \bar{X} \\ \hat{m}^2\hat{p}^2 + \hat{m}\hat{p}(1-\hat{p}) = \frac{1}{n}\sum X_i^2 \end{cases}$

$$\Rightarrow \hat{p} = 1 - (M_2 - M_1^2)/M_1, \hat{m} = M_1/\hat{p}.$$

(Bernoulli population vs binomial population.)

$$\begin{cases} \hat{m}\hat{p} = \bar{X} \\ (\hat{m}\hat{p})^2 + \hat{m}\hat{p}(1-\hat{p}) = \frac{1}{n}\sum X_i^2 \\ \bar{X}^2 + \bar{X}(1-\hat{p}) = M_2 \end{cases}$$

$$\boxed{\begin{cases} \hat{p} = 1 - \frac{M_2 - \bar{X}^2}{\bar{X}} \\ \hat{m} = \bar{X}/\hat{p} \end{cases}}$$

2^2

Given a sample

$$2, 5, 6, 3$$

of size ~~$n = 4$ from a binomial~~ population. $\bar{x} = 4$, $S^2 =$

$M_2 = \frac{1}{4}(4^2 + 0^2 + 1^2 + 6^2) = 18.5.$ $M_2 = \bar{x}^2 + \frac{n-1}{n} S^2$

$\hat{p} = 1 - (18.5 - 4^2)/4 = 0.375,$ $\hat{m} = 4/0.375 = 10.67 \approx 11.$

Correction: $M_2 = \frac{1}{4}(2^2 + 5^2 + 6^2 + 3^2) = 18.5.$

$$M_2 = \frac{1}{n}\sum x_i^2 = \frac{1}{4}(2^2 + 5^2 + 6^2 + 3^2) = 18.5$$

Example (Gamma) Let X be gamma with shape α and scale β.

$$E(X) = \alpha\beta, \mathrm{Var}(X) = \alpha\beta^2.$$

Set sample moments to theoretical moments:

$$\begin{cases} \alpha\beta = \bar{X} \\ \alpha\beta^2 + (\alpha\beta)^2 = M_2 \end{cases}$$

$$\hat{\beta} = \frac{M_2 - \bar{X}^2}{\bar{X}} = \frac{(n-1)S^2}{n\bar{X}}, \hat{\alpha} = \frac{\bar{X}}{\hat{\beta}} = \frac{n\bar{X}^2}{(n-1)S^2}.$$

scale β. $f(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, x > 0$

rate $= \frac{1}{\text{scale}} = \frac{1}{\beta} = \lambda.$

$$f(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, x > 0$$

$$E(X) = \frac{\alpha}{\lambda}, V(X) = \frac{\alpha}{\lambda^2}.$$

$$\begin{cases} E(X) = \alpha\beta = \bar{X} \\ E(X^2) = \mu^2 + \sigma^2 = (\alpha\beta)^2 + \alpha\beta^2 = M_2 \end{cases}$$

$$\bar{X}^2 + \bar{X}\beta = M_2$$

$$\Rightarrow \begin{cases} \hat{\beta} = \frac{M_2 - \bar{X}^2}{\bar{X}} \\ \hat{\alpha} = \frac{\bar{X}}{\hat{\beta}} \end{cases}$$

For a normal population with mean μ and variance σ^2, we set sample moments to theoretical moments:

$$\begin{cases} \mu = \bar{X} \\ \mu^2 + \sigma^2 = M_2 \end{cases}$$

$$\hat{\mu} = \bar{X}, \hat{\sigma}^2 = \frac{n-1}{n} S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

(The method of moments estimator of σ^2 is slightly different from the usual unbiased S^2.)

$$\hat{\sigma}^2 = M_2 - \bar{X}^2$$
$$= (\bar{X}^2 + \frac{n-1}{n} S^2) - \bar{X}^2 = \frac{n-1}{n} S^2$$
$$= \frac{1}{n}\sum (X_i - \bar{X})^2$$