1/18/22

Tuesday, January 18, 2022 7:34 PM

# $\boxed{\text{Statistics}}$

We will typically be in a situation where we find some

$$\boxed{X_1, X_2, \dots, X_n \text{ iid with common pdf } f_{X_i}(x_i)}$$

We should think of this as a "repeated experiment"

Some Notation: $f_X(x;\theta)$

pdf or pmf.

$\theta$ is an unknown parameter.

there could also be multiple unknown parameters.

we say $\theta \in \Theta$, where $\Theta$ encompasses all possible values of our unknown parameter $\theta$.

remember that A statistic is a function of the data.

eg. $\boxed{h(X_1, X_2, \dots, X_n)}$

* notice we do not need to know $\theta$.

Examples: $N(\mu = 2, \sigma^2)$ where $\sigma$ is unknown

here, we say $\sum (x_i - \mu)^2$ is a statistic because we know $\mu$ & $x_i$'s.

* if $\mu$ was unknown then this would not be a statistic

A $\boxed{\text{Sampling Distribution}}$ $X_1, X_2, \dots, X_n$ iid $\sim f_{X_i}(x_i, \theta)$

$Y = h(X_1, \dots, X_n)$ is a statistic

Then the distribution of $Y$, $f_Y(y)$ is a sampling distribution.

We can reference this 'more precisely' as $f_Y(y;\theta)$

We can reference this 'more precisely' as $f_Y(y;\vartheta)$

Consider our usual situation...

What we may want to do is estimate our paramater $\vartheta$

We say a statistic, $y = h(\text{data})$, whose purpos is to estimate $\vartheta$, is refered to as an ESTIMATOR

Example: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ is an estimator of $\mu$.

"sample mean" | | | | "population mean"

$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is an estimator of $\sigma^2$

"sample variance" | | | | "population variance"

Assume we are working wit $X_1, \dots, X_n$ iid $\sim N(\mu, \sigma^2)$

↳ We know that $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

↳ We also know that $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

↳ so we then know the pdf of $S^2$

Some Facts: $\mathbb{E}[\chi^2_{n-1}] = n-1$ & $\mathbb{V}[\chi^2_{n-1}] = 2(n-1)$

Thus $\mathbb{E}[S^2] = \sigma^2$ & $\mathbb{V}[S^2] = \frac{2\sigma^4}{n-1}$

As $\frac{(n-1)}{\sigma^2}\mathbb{E}[S^2] = n-1$ As $\frac{(n-1)^2}{\sigma^4}\mathbb{V}[S^2] = 2(n-1)$

$\mathbb{E}[\frac{(n-1)S^2}{\sigma^2}]$ $\mathbb{V}[\frac{(n-1)S^2}{\sigma^2}]$

We can pull Constants out of our Expected Value & Our Variance ★

Some Additional Notation:

We say that $T$ is an <u>unbiased estimator</u> of $\theta$ if $\mathbb{E}[T] = \theta \ \forall \theta \in \Theta$

For example $\bar{X}$, regardless of $f_X(X, \theta)$, so long as $\mathbb{E}[X]$ exists,

then $\mathbb{E}[\bar{X}] = \mathbb{E}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{n\mu}{n} = \mu$.

$\therefore$ $\bar{X}$ is an unbiased estimator of $\mu$.

Likewise, so long as $\mathbb{E}[X^2]$ exists, then

$\mathbb{E}[S^2] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \sigma^2$ & then

$S^2$ is an unbiased estimator of $\sigma^2$.

Note: in general, $S$ is <u>not</u> an unbiased estimator of $\sigma$.

We will often want to find unbiased estimators.

Ex. Given $X_1, \dots, X_n$ are iid $\sim Exp[\beta]$ $\therefore \left\{ f_{X_i}(x_i) = \frac{1}{\beta} e^{-\frac{x}{\beta}}, \ \mathbb{E}[X_i] = \beta, \ \mathbb{V}[X_i] = \beta^2 \right\}$

- We would like to find an unbiased estimator of $\frac{1}{\beta}$
- We know that $\bar{X}$ is a good estimator of $\mu$
  - thus $\bar{X}$ is a good estimator of $\beta$.
- Naturally, let's try to use $T = \frac{1}{\bar{X}}$ as an estimator of $\frac{1}{\beta}$.

Notice $T = \frac{1}{\bar{X}} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{n}{\Sigma}$ ($\Sigma$ as shorthand for $\sum_{i=1}^{n} x_i$)

So let's find the pdf of $\frac{1}{\Sigma}$

Recall: $X_1, \dots, X_n$ are iid $Exp(\beta)$ so $\Sigma$ is $Gamma(\alpha = n, \beta = \beta)$

The pdf of $Gamma(\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$, $x > 0$

Recall: If $\alpha$ is a positive integer, then, $\Gamma(\alpha) = (\alpha - 1)!$

Also: $\int_0^\infty x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \Gamma(\alpha)\beta^\alpha = (\alpha-1)!\beta^\alpha$

Now we can try to find $\mathbb{E}[\frac{1}{\Sigma}]$:

$$\mathbb{E}[\tfrac{1}{\Sigma}] = \int_0^\infty \left(\frac{1}{x}\right) \cdot \frac{1}{\Gamma(n)\beta^n} x^{n-1} e^{-\frac{x}{\beta}} dx$$

$$= \int_0^\infty \frac{1}{\Gamma(n)\beta^n} x^{n-2} e^{-\frac{x}{\beta}} dx$$

$$= \frac{1}{\Gamma(n)\beta^n} \cdot \int_0^\infty x^{n-2} e^{-\frac{x}{\beta}} dx$$

$$= \frac{1}{\Gamma(n)\beta^n} \cdot \Gamma(n-1)\beta^{n-1}$$

$$\boxed{\mathbb{E}\left[\frac{1}{\sum_{i=1}^n X_i}\right]} = \boxed{\mathbb{E}[\tfrac{1}{\Sigma}]} = \frac{(n-2)!}{(n-1)!} \cdot \frac{\beta^{n-1}}{\beta^n} = \boxed{\frac{1}{(n-1)\beta}}$$

Note: $(n-1) \cdot \mathbb{E}[\frac{1}{\Sigma}] = \frac{1}{\beta}$

So our question is: is $\frac{1}{\bar{X}} = \frac{n}{\Sigma}$ an unbiased estimator of $\frac{1}{\beta}$?

$\bar{X} = \frac{\Sigma}{n}$ well: $\mathbb{E}[\frac{n}{\Sigma}] = \frac{n}{(n-1)\beta} = \mathbb{E}[\frac{1}{\bar{X}}]$

this is not as we want $\frac{1}{\beta}$

Thus, a true unbiased estimator of $\frac{1}{\beta}$ is $\frac{n-1}{n} \cdot \frac{1}{\bar{X}} = T$

Note: $T = \frac{n-1}{n} \cdot \frac{n}{\Sigma} = \frac{n-1}{\Sigma}$

$$\boxed{\text{thus we use } T = \frac{n-1}{\sum_{i=1}^{n} X_i} \text{ to estimate } \frac{1}{\beta}}$$

This is A Common Techinique We will be using!

* Guess the Estimator
* Figure out a fudge factor
* Construct the Estimator