Welcome Today we'll discuss Worksheet 13_Part2

Section 6.2: Method of Moments (MOM) Estimates

- Let X be a random variable with pdf $f(x; \theta_1, \theta_2, \ldots \theta_k)$ that depends on parameters $\theta_1, \theta_2, \ldots, \theta_k$.

- If we independently pick a random variables $X_1, X_2, \ldots X_n$ from population X, we can determine what values of $\theta_1, \theta_2, \ldots, \theta_k$ that best fit the data in the following sense:

  1. The mean of the population X equals the sample mean.
  2. The variance of the population equals the variance of the sample.
  3. The skewness of the population equals the skewness of the sample.
  4. ... and so on.

Goal: Our pop should have same properties as observed in sample

- We find values of the parameters so the properties of random variable X are equal to corresponding properties of our sample.

Let X be a random variable with pdf $f(x)$. For a positive integer k, **the kth theoretical moment of X** is

$$\mu_k = E[X^k] = \int_{-\infty}^{\infty} x^k f(x)\, dx \quad \text{or} \quad \mu_k = E[X^k] = \sum_X x^k f(x),$$

Depend on parameters

- The mean $\mu = E[X]$ is the first moment.
- The variance is related to the second moment $\mu_2 = E[X^2]$. $Var(X) = E(X^2) - (E(X))^2$
- The skewness is related the third moment $\mu_3 = E[X^3]$.
- The kurtosis (how "peaky" or flat the distribution is) is related to $\mu_4 = E[X^4]$.

For a sample we called the corresponding properties **sample moments** denoted by M_k.

Values we calculate

Section 6.2: Method of Moments Estimate

1. Let X be a random variable with pdf $f(x;\lambda,\delta) = \boxed{\lambda e^{-\lambda(x-\delta)}}$ for $\underline{x > \delta}$ with parameters $\lambda, \delta > 0$. Find the first and second theoretical moments of X.

$$\mu_1 = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{\delta}^{\infty} x\left(\lambda e^{-\lambda(x-\delta)}\right)dx = \delta + \frac{1}{\lambda}$$

$$\mu_2 = E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_{\delta}^{\infty} x^2\left(\lambda e^{-\lambda(x-\delta)}\right)dx = \left(\delta + \frac{1}{\lambda}\right)^2 + \frac{1}{\lambda^2}$$

2. Let $X_1 = 3$, $X_2 = 4$, $X_3 = 5$, and $X_4 = 8$ be a random sample from a random variable X with pdf $f(x;\lambda,\delta)$. Find the first and second sample moments.

$$M_1 = \frac{3+4+5+8}{4} = 5$$

$$M_2 = \frac{3^2+4^2+5^2+8^2}{4} = 28.5$$

Let X be a random variable with pdf $f(x;\theta_1,\theta_2,\ldots,\theta_k)$ and let X_1, X_2, …, X_n be a random sample.

- The theoretical moments μ_k are functions of the k parameters $\theta_1,\theta_2,\ldots,\theta_k$.
- The sample moments M_k are values we calculate based on the sample.
- The method of moments (MOM) estimate is obtained by solving the system:

$$\begin{aligned}
\mu_1 &= \int_{-\infty}^{\infty} x f(x)\,dx = \frac{1}{n}\sum_{i=1}^{n} X_i = M_1\\
\mu_2 &= \int_{-\infty}^{\infty} x^2 f(x)\,dx = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = M_2\\
&\vdots\\
\mu_k &= \int_{-\infty}^{\infty} x^k f(x)\,dx = \frac{1}{n}\sum_{i=1}^{n} X_i^k = M_k
\end{aligned}$$

If X is a discrete random variable, change the integrals to summations.

Practice

3. Let X be a random variable with pdf $f(x; \lambda, \delta) = \lambda e^{-\lambda(x-\delta)}$ for $x > \delta$ with parameters $\lambda, \delta > 0$. If $X_1 = 3$, $X_2 = 4$, $X_3 = 5$, and $X_4 = 8$ is a random sample picked from random variable X, find the first and second sample moments.

① Look at how many parameters are in your model

2 - λ and δ

② Calculate as many sample moments as there are parameters (Question 2)

③ Calculate as many theoretical moments as there are parameters (Question 1)

④ set sample and theoretical moments equal, and solve for parameters

4. Let $X_1 = 1, X_2 = 3, X_3 = 7, X_4 = 10$ be four numbers picked at random from the uniform distribution on $[\alpha, \beta]$. Find the MoM estimates of α and β.

Theoretical	Sample
$\mu_1 = \delta + \frac{1}{\lambda}$	$m_1 = \frac{3+4+5+8}{4} = 5$
$\mu_2 = \left(\delta + \frac{1}{\lambda}\right)^2 + \frac{1}{\lambda^2}$	$m_2 = \frac{3^2+4^2+5^2+8^2}{4} = 28.5$

$$\begin{cases} \delta + \frac{1}{\lambda} = 5 & (1^{st} \text{ moment}) \\ \left(\delta + \frac{1}{\lambda}\right)^2 + \frac{1}{\lambda^2} = 28.5 & (2^{nd} \text{ moment}) \end{cases}$$

$$(5)^2 + \frac{1}{\lambda^2} = 28.5$$

$$\frac{1}{\lambda^2} = 28.5 - 25$$

$$\lambda^2 = \frac{1}{3.5}$$

$$\boxed{\hat{\lambda}_{mom} \approx 0.535}$$

$$\delta + \frac{1}{0.535} = 5$$

$$\boxed{\hat{\delta}_{mom} = 3.131}$$

4. Let $X_1 = 1, X_2 = 3, X_3 = 7, X_4 = 10$ be four numbers picked at random from the uniform distribution on $[\alpha, \beta]$. Find the MoM estimates of α and β.

① Look at how many parameters are in your model

2 – α and β

② Calculate as many sample moments as there are parameters

$$m_1 = \frac{1+3+7+10}{4} = 5.25 \qquad m_2 = \frac{1^2+3^2+7^2+10^2}{4} = 39.75$$

③ Calculate as many theoretical moments as there are parameters

Recall: $\boxed{E(X) = \frac{\alpha+\beta}{2}}$

$$Var(X) = \frac{(\beta-\alpha)^2}{12} = E(X^2) - (E(X))^2$$

$$\boxed{E(X^2) = \frac{(\beta-\alpha)^2}{12} + \left(\frac{\alpha+\beta}{2}\right)^2}$$

④ Set sample and theoretical moments equal, and solve for parameters

$$\frac{\alpha+\beta}{2} = 5.25$$

$$\frac{(\beta-\alpha)^2}{12} + \left(\frac{\alpha+\beta}{2}\right)^2 = 39.75$$

$\boxed{\text{Solve}}$

$$\frac{\alpha+\beta}{2} = 5.25$$

$$\frac{(\beta-\alpha)^2}{12} + \left(\frac{\alpha+\beta}{2}\right)^2 = 39.75$$

$$\left(\frac{\beta-\alpha}{12}\right)^2 + (5.25)^2 = 39.75$$

$$\beta - \alpha = 12.093$$

$$+\ \beta + \alpha = 10.5$$

$$2\beta = 22.593$$

$$\boxed{\hat{\beta}_{mom} = 11.297} \Longrightarrow \alpha + 11.297 = 10.5$$

$$\boxed{\hat{\alpha}_{mom} = -0.797}$$

$$X_1 = 1,\ X_2 = 3,\ X_3 = 7,\ X_4 = 10$$

Section 6.3.1: Unbiasedness

Which method is best? We will wrap up Chapter 6 by looking at some properties we can use to gauge estimates, such as unbiasedness, efficiency, and Mean Square Error.

Bias: We like an estimator to be, on average, equal to the parameter it is estimating: $\boxed{E[\hat{\theta}] - \theta} = 0$.

- Sample mean is unbiased estimator of μ.
- Sample proportion is an unbiased estimator of p.

5. Show that the MLE estimate for the variance for $X \sim N(\mu, \sigma^2)$,

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

☆ Optional

is a biased estimate for σ^2

If $X \sim N(\mu, \sigma)$. $X_1, X_2, \ldots, X_n$

MLE:

$\hat{\mu}_{MLE} = \bar{X} = \frac{\sum X_i}{n}$

$\hat{\sigma}_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

$$E(\hat{\mu}_{MLE}) - \mu = E\left(\frac{\sum X_i}{n}\right) - \mu$$
$$= \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) - \mu$$
$$= \frac{1}{n}(\mu_X \cdot n) - \mu = 0$$

$E(\hat{\sigma}^2_{MLE}) - \sigma^2 = 0$??

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

$$E(\hat{\theta}^2) = \frac{1}{n} E\left(\sum (x_i - \bar{x})^2\right)$$

$$\vdots$$

$$= \frac{n-1}{n}\theta^2 \neq \theta^2$$

Optional

As $n \to \infty$ $E(\hat{\theta}^2) \to \theta^2$

Asymptotically unbiased

$$\hat{S}^2 = \frac{n\,\hat{\theta}^2}{n-1} \rightarrow E(\hat{S}^2) = \frac{n}{n-1}E(\hat{\theta}^2)$$

$$= \frac{n}{n-1}\left(\frac{n-1}{n}\theta^2\right)$$

$$= \theta^2$$

$$\hat{S}^2 = \frac{n}{n-1}\left(\frac{\sum (x_i - \bar{x})^2}{n}\right)$$

$$\boxed{\hat{S}^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}}$$

is an unbiased estimator for variance

$E(Y) = \mu$?

$$Y = \tfrac{1}{6}X_1 + \tfrac{1}{3}X_2 + \tfrac{1}{2}X_3$$

$$E(Y) = \tfrac{1}{6}E(X_1) + \tfrac{1}{3}E(X_2) + \tfrac{1}{2}E(X_3)$$

$$= \tfrac{1}{6}\mu_x + \tfrac{1}{3}\mu_x + \tfrac{1}{2}\mu_x = \mu_x \checkmark$$

Section 6.3.2: Efficiency

Let X_1, X_2, X_3 be independent random variables from an identical distribution with mean and variance μ and σ^2, respectively.

On HW (not optional)

- We have shown $\bar{X} = \dfrac{X_1 + X_2 + X_3}{3}$ is an unbiased estimator of μ.
- The weighted mean $Y = \dfrac{1}{6}X_1 + \dfrac{1}{3}X_2 + \dfrac{1}{2}X_3$ is also unbiased.
- Is one better than the other?

6. We have two unbiased estimators of μ given below. Which estimator has less variability?

$$\bar{X} = \frac{X_1 + X_2 + X_3}{3} \quad \text{and} \quad Y = \frac{1}{6}X_1 + \frac{1}{3}X_2 + \frac{1}{2}X_3.$$

If $E(\hat{\mu}) = \mu$ unbiased, great!

A little bias might not be so bad if $Var(\hat{\mu})$ is small (efficient)

$$\text{Var}[\bar{X}] = \text{Var}\left[\frac{X_1 + X_2 + X_3}{3}\right] = \frac{1}{3^2}(3\sigma^2) = \frac{\sigma^2}{3}.$$

Less variance

$\frac{1}{3} < \frac{7}{18}$

$$\text{Var}[Y] = \text{Var}\left[\frac{1}{6}X_1 + \frac{1}{3}X_2 + \frac{1}{2}X_3\right] = \left(\frac{1}{36} + \frac{1}{9} + \frac{1}{4}\right)(3\sigma^2) = \frac{7}{18}\sigma^2.$$

If θ_1 and θ_2 are both unbiased estimators of θ, then θ_1 is said to be more efficient than θ_2 if $\text{Var}[\theta_1] < \text{Var}[\theta_2]$.

Section 6.3.3: Mean Square Error

Optional

The Mean Square Error (MSE) of an estimator measures the average squared distance between the estimator and the parameter:

$$\text{MSE}[\hat{\theta}] = \text{E}[(\hat{\theta} - \theta)^2].$$

Proposition 6.3.3 shows that $\text{MSE}[\hat{\theta}] = \text{Var}[\hat{\theta}] + (\text{Bias}[\hat{\theta}])^2$.

- MSE is a criterion that combines bias and variance.
- If two estimators are unbiased, one is more efficient than the other if and only if its MSE is smaller.
- In general, we are often faced with a trade-off between variability and bias.

7. Let $X \sim \text{Binom}(n, p)$ with n known and parameter p unknown.

   (a) Find the variance and MSE for if we use the sample proportion $\hat{p}_1 = \frac{X}{n}$ as an estimate for p.

   $$E(\hat{p}_1) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \left(\frac{1}{n}\right)(np) = p$$

   (b) If we add two more trials to the sample, and assume one is a failure and the other a success, then we can define a second estimator for p

   $$\underline{\hat{p}_2 = \frac{X+1}{n+2}.}$$

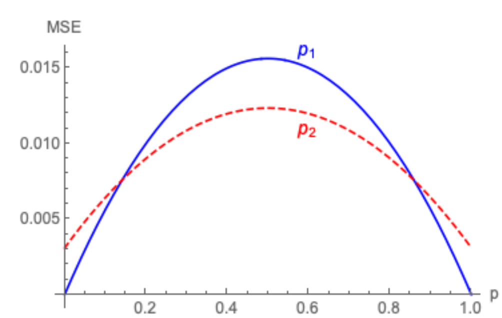
   Added two attempts to our sample, one is a success.

   (c) Is $\hat{p}_2$ is a biased or unbiased estimator for p?

   $$E(\hat{p}_2) = \frac{1}{n+2}E(X+1) = \frac{1}{n+2}\left(E(X) + E(1)\right) = \frac{np+1}{n+2}$$

   (d) Find the Var $[\hat{p}_2]$ and MSE $[\hat{p}_2]$.

$\boxed{\text{Optional}}$



- We have looked at two more estimators for unknown population parameters: MLE and MoM.
- We looked at different criteria to compare different estimators: Bias, Efficiency, and MSE.
- The text discusses more properties at the end of Section 6.3 that are more technical.