- Note:

$$E\left[g(X_i) - E(g(X_i))\right]^2 = Var\, g(X_i) = Var\, g(X_1) \tag{7}$$

- and

$$E\left(\left[g(X_i) - E(g(X_i))\right]\left[g(X_j) - E(g(X_j))\right]\right) = Cov\left[g(X_i),\ g(X_j)\right] \tag{8}$$

- So

$$\begin{aligned} Var\left(\sum_{i=1}^{n} g(X_i)\right) &= E\left[\sum_{i=1}^{n} \left(g(X_i) - E\left(g(X_i)\right)\right)\right]^2 \\ = \quad & E\left[\sum_{i=1}^{n} \left(g(X_i) - E\left(g(X_i)\right)\right)^2\right] \\ & + E\left[2\sum_{i=1}^{n}\sum_{j=i}^{n} \left(g(X_i) - E\left(g(X_i)\right)\right)\left(g(X_j) - E\left(g(X_j)\right)\right)\right] \end{aligned}$$

- and

$$
\begin{aligned}
& Var\left(\sum_{i=1}^{n} g(X_i)\right) \\
= \quad & \sum_{i=1}^{n} E\left[\left(g(X_i) - E\left(g(X_i)\right)\right)^2\right] \\
& +2\sum_{i=1}^{n}\sum_{j=i}^{n} E\left[\left(g(X_i) - E\left(g(X_i)\right)\right)\left(g(X_j) - E\left(g(X_j)\right)\right)\right]
\end{aligned}
$$

- Recall that the X_1 and X_j in the sample are independent, so each of the covariances is zero.
- Now rewrite:

$$
\begin{aligned}
& Var\left(\sum_{i=1}^{n} g(X_i)\right) \\
= \; & \sum_{i=1}^{n} E\left[(g(X_i) - E(g(X_i))^2\right] \\
& +2\sum_{i=1}^{n}\sum_{j=i}^{n} E\left[(g(X_i) - E(g(X_i))(g(X_j) - E(g(X_j))\right] \\
= \; & \sum_{i=1}^{n} Var\, g(X_i) = \sum_{i=1}^{n} Var\, g(X_1) = n \;\; Var\, g(X_1)
\end{aligned}
$$

Distribution of Sample Statistics

Theorem

Let $X_1, X_2, \cdots X_n$ *be a random sample from a population with mean* μ *and variance* $\sigma^2 < \infty$. *Then*

a) $E\bar{X} = \mu$

b) $Var\bar{X} = \frac{\sigma^2}{n}$

c) $ES^2 = \sigma^2$

- **Proof of part a).** In theorem 2 let

$$g(X) = g(X_i) = \frac{X_i}{n}$$

- Then

$$Eg(X_i) = \frac{\mu}{n}$$

- Then we can write

$$\begin{aligned} E\bar{X} &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = E\left(\sum_{i=1}^{n}\frac{X_i}{n}\right) \\ &= E\left(\sum_{i=1}^{n} g(X_i)\right) = nE\left(g(X_i)\right) = n\frac{\mu}{n} = \mu \end{aligned}$$

- **Proof of part b).** In theorem 2 let

$$g(X) = g(X_i) = \frac{X_i}{n}$$

- Then

$$Var\left[g(X_i)\right] = \frac{\sigma^2}{n^2}$$

- Then we can write

$$\begin{aligned} Var\bar{X} &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = Var\left(\sum_{i=1}^{n}\frac{X_i}{n}\right) \qquad (9) \\ &= Var\left(\sum_{i=1}^{n} g(X_i)\right) = n\left(Varg(X_1)\right) = n\frac{\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

- **Proof of part c).**
- Write S^2 as

$$
\begin{aligned}
ES^2 &= E\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right) = E\left(\frac{1}{n-1}\left[\sum_{i=1}^{n}X_i^2 - n\bar{X}^2\right]\right) \\
&= \frac{1}{n-1}\left(nEX_1^2 - nE\bar{X}^2\right) \\
&= \frac{1}{n-1}\left(n(\sigma^2+\mu^2) - n\left(\frac{\sigma^2}{n}+\mu^2\right)\right) = \sigma^2
\end{aligned}
$$

- The last line follows from

$$
\begin{aligned}
Var\ X &= \sigma_X^2 = EX^2 - (EX)^2 \\
&= EX^2 - \mu_X^2 \\
\Rightarrow E\ X^2 &= \sigma_X^2 + \mu_X^2
\end{aligned}
$$

Unbiased Statistics

Definition

A statistic $T(X)$ is unbiased for the parameter θ if $ET(X) = \theta$.

- Note that then in a random sample:
  - $\bar{X}$ is an unbiased statistic for μ
  - S^2 is an unbiased statistic for σ^2

- Let $Y_1, Y_2, \cdots Y_n$ denote a random sample from a population characterized by the parameters $\theta_1, \theta_2, \cdots \theta_k$.
- Y has pdf $f(\ \cdot; \theta_1, \theta_2, \cdots \theta_k)$.
- First estimation method: Method of Moments:
- Recall:
    - The r^{th} moment of Y is defined as

$$\mu'_r = E(Y^r)$$

    - The rth sample moment is defined as

$$\hat{\mu}'_r = \bar{x}^r_n = \frac{1}{n}\sum_{i=1}^{n} y_i^r$$

- In general μ'_r is a known function of $\theta_1, \theta_2, \cdots \theta_k$, say $\mu'_r = g_r(\theta_1, \theta_2, \cdots \theta_k)$.
- Form the K equations, equating population and sample moments:

$$\begin{aligned}
\mu'_1 &= g_1(\theta_1, \theta_2, \cdots \theta_k) = \hat{\mu}'_1 = \frac{1}{n}\sum_{i=1}^{n} y_i \\
\mu'_2 &= g_2(\theta_1, \theta_2, \cdots \theta_k) = \hat{\mu}'_2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 \\
&\vdots \\
\mu'_K &= g_K(\theta_1, \theta_2, \cdots \theta_k) = \hat{\mu}'_K = \frac{1}{n}\sum_{i=1}^{n} y_i^K
\end{aligned}$$

- Solve these K equations for $\theta_1, \theta_2, \cdots \theta_k$
- The result is the estimators, denoted $\hat{\theta}_1, \hat{\theta_2}, \cdots \hat{\theta_K}$.

- Let $Y_1, Y_2, \cdots Y_n$ denote a random sample from the pdf:

$$f(y) = \begin{cases} (p+1)\, y^p & 0 \leq y \leq 1 \\ 0 & \textit{otherwise} \end{cases}$$

- The first moment of Y is:

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y\, f(y)\, dy = \int_0^1 y\,(p+1)\, y^p\, dy = \int_0^1 y^{p+1}\,(p+1)\, dy \\ &= \left. \frac{y^{p+2}\,(p+1)}{(p+2)} \right|_0^1 = \frac{p+1}{p+2} \end{aligned}$$

- Set $E(Y)$ equal to the first sample moment and solve for p:

$$\begin{aligned}
\frac{p+1}{p+2} &= \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y} \\
\Rightarrow p+1 &= (p+2)\,\bar{y} = p\bar{y} + 2\bar{y} \\
\Rightarrow p - p\bar{y} &= 2\,\bar{y} - 1 \\
\Rightarrow p(1-\bar{y}) &= 2\,\bar{y} - 1 \\
\Rightarrow \hat{p} &= \frac{2\,\bar{y} - 1}{1-\bar{y}}
\end{aligned}$$

- Let $Y_1, Y_2, \cdots Y_n$ denote a random sample from a normal distribution with mean μ and variance σ^2.
- Recall that

$$\begin{aligned} \mu_1' &= E(Y) = \mu \\ \mu_2' &= E(Y^2) = \sigma^2 + E^2[Y] = \sigma^2 + \mu^2 \end{aligned}$$

- Now set the first population moment equal to its sample analogue:

$$\begin{aligned} \mu &= \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y} \qquad (10) \\ \Rightarrow \hat{\mu} &= \bar{y} \end{aligned}$$

- Now set the second population moment equal to its sample analogue

$$
\begin{aligned}
\sigma^2 + \mu^2 &= \frac{1}{n}\sum_{i=1}^{n} y_i^2 \\
\Rightarrow \sigma^2 &= \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \mu^2 \\
\Rightarrow \sigma &= \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2 - \mu^2}
\end{aligned} \tag{11}
$$

- Plug in $\hat{\mu} = \bar{y}$:

$$
\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}^2} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n}}
$$

- Different estimate of standard deviation.

- Let $X_1, X_2, \cdots X_n$ denote a random sample from a gamma distribution with parameters α and β.
- The pdf is

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{\frac{-x}{\beta}} & 0 \leq x < \infty \\ 0 & otherwise \end{cases}$$

- Find the first moment:

$$E(X) = \int_0^\infty x \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{\frac{-x}{\beta}} dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty x^{(1+\alpha)-1} e^{\frac{-x}{\beta}} dx$$

- Multiply and divide by $\beta^{1+\alpha} \ \Gamma(1+\alpha)$

$$\begin{aligned} E(X) &= \frac{\beta^{1+\alpha} \ \Gamma(1+\alpha)}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty \frac{1}{\beta^{1+\alpha} \ \Gamma(1+\alpha)} x^{(1+\alpha)-1} e^{\frac{-x}{\beta}} dx \\ &= \frac{\beta^{1+\alpha} \ \Gamma(1+\alpha)}{\beta^\alpha \Gamma(\alpha)} \cdot 1 = \frac{\beta \ \Gamma(1+\alpha)}{\Gamma(\alpha)} \end{aligned}$$

- Recall

$$\Gamma(v+1) = v\Gamma(v)$$

- Using this

$$E(X) = \frac{\beta\,\Gamma(1+\alpha)}{\Gamma(\alpha)} = \frac{\beta\,\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \beta\,\alpha$$

- Now find $E(X^2)$:

$$E(X^2) = \int_0^\infty x^2 \frac{1}{\beta^\alpha\Gamma(\alpha)}\, x^{\alpha-1}\, e^{\frac{-x}{\beta}}\, dx = \frac{1}{\beta^\alpha\Gamma(\alpha)} \int_0^\infty x^{(2+\alpha)-1}\, e^{\frac{-x}{\beta}}\, dx$$

- Multiply and divide by $\beta^{2+\alpha}\ \ \Gamma(2+\alpha)$

$$
\begin{aligned}
E(X^2) &= \frac{\beta^{2+\alpha}\ \Gamma(2+\alpha)}{\beta^{\alpha}\Gamma(\alpha)} \int_0^\infty \frac{1}{\beta^{2+\alpha}\ \Gamma(2+\alpha)} x^{(2+\alpha)-1}\, e^{\frac{-x}{\beta}}\, dx \\
&= \frac{\beta^{2+\alpha}\,\Gamma(2+\alpha)}{\beta^{\alpha}\Gamma(\alpha)} = \frac{\beta^2(\alpha+1)\,\Gamma(1+\alpha)}{\Gamma(\alpha)} = \frac{\beta^2\alpha(\alpha+1)\,\Gamma(\alpha)}{\Gamma(\alpha)} \\
&= \beta^2\alpha(\alpha+1)
\end{aligned}
$$

- Now set the first population moment equal to the sample analogue:

$$
\begin{aligned}
\beta\,\alpha &= \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \\
\Rightarrow\ \hat{\alpha} &= \frac{\bar{x}}{\beta}
\end{aligned}
$$

- Now set the second population moment equal to its sample analogue

$$\beta^2 \alpha(\alpha + 1) = \frac{1}{n} \sum_{i=1}^{n} x_i^2$$

$$\Rightarrow \beta^2 = \frac{\sum_{i=1}^{n} x_i^2}{n\alpha(\alpha + 1)}$$

$$\Rightarrow \beta^2 = \frac{\sum_{i=1}^{n} x_i^2}{n\left(\frac{\bar{x}}{\beta}\right)\left(\left(\frac{\bar{x}}{\beta}\right) + 1\right)} = \frac{\sum_{i=1}^{n} x_i^2}{\left(\frac{n\bar{x}^2}{\beta^2}\right) + \left(\frac{n\bar{x}}{\beta}\right)}$$

$$\Rightarrow n\bar{x}^2 + n\bar{x}\beta = \sum_{i=1}^{n} x_i^2$$

$$n\,\bar{x}^2 + n\,\bar{x}\,\beta = \sum_{i=1}^{n} x_i^2$$

$$\Rightarrow n\,\bar{x}\;\beta = \sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2$$

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

- Thus, we estimated both parameters.

- Let $Y_1, Y_2, \cdots Y_n$ denote a random sample from an unknown distribution with parameters β and σ.
- We know the following about the distribution of Y:

$$\begin{aligned} Y &= \beta + u \\ E(u) &= 0 \\ E(u^2) &= Var(u) = \sigma^2 \end{aligned}$$

- Consider estimators for β and σ^2:
  - Note that we know the moments of u, so we need the sample moments of u, not y.
  - In a given sample

$$\begin{aligned} y_i &= \hat{\beta} + \hat{u}_i \qquad (12) \\ \Rightarrow \quad \hat{u}_i &= y_i - \hat{\beta} \end{aligned}$$

- The sample moments are then as follows

$$\text{First sample moment} = \frac{\sum_{i=1}^{n} \hat{u}_i}{n}$$

$$\text{Second sample moment} = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n}$$

- or, substituting from (12)

$$\text{First sample moment} = \frac{\sum_{i=1}^{n} \left(y_i - \hat{\beta}\right)}{n}$$

$$\text{Second sample moment} = \frac{\sum_{i=1}^{n} \left(y_i - \hat{\beta}\right)^2}{n}$$

- Set the first sample moment equal to the first population moment:

$$\frac{\sum_{i=1}^{n}\left(y_i - \hat{\beta}\right)}{n} = 0$$

$$\Rightarrow \sum_{i=1}^{n} y_i = n\hat{\beta}$$

$$\Rightarrow \hat{\beta} = \frac{\sum_{i=1}^{n} y_i}{n} = \bar{y}$$

- Set the second sample moment equal to the second population moment:

$$\frac{\sum_{i=1}^{n}\left(y_i - \hat{\beta}\right)^2}{n} = \sigma^2$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(y_i - \hat{\beta}\right)^2}{n} = \frac{\sum_{i=1}^{n}\left(y_i - \hat{\beta}\right)^2}{n}$$

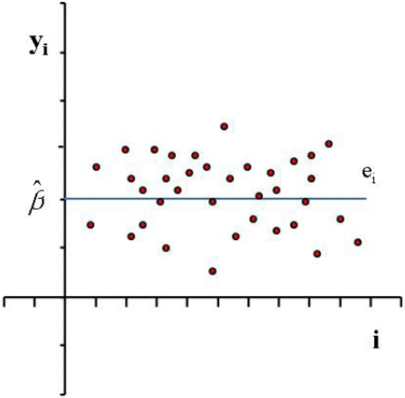
- LS estimation: Example with one parameter
- Consider Y_i such that for all i

$$\begin{aligned} Y_i &= \beta + \epsilon_i \\ E(\epsilon_i) &= 0 \\ Var(\epsilon_i) &= \sigma^2 \end{aligned}$$

- Then y_i is drawn from a population with mean of β and a variance of σ^2.
- The least squares estimator of β is obtained by minimizing the sum of squares defined by

$$SSE = \sum_{i=1}^{n} (y_i - \hat{\beta})^2 \tag{13}$$

- Intuition:



- Find $\hat{\beta}$

$$
\begin{aligned}
\frac{\partial SSE}{\partial \beta} &= 2\sum_{i=1}^{n}(y_i - \hat{\beta})(-1) = 0 \\
\Rightarrow \quad \hat{\beta} &= \frac{1}{n}\sum_{i-1}^{n} y_i = \bar{y}
\end{aligned}
$$

- The distance between the elements of the random sample and "predicted" values are minimized.

- LS estimation: Example with two parameters:
- Consider the model

$$\begin{aligned} y_t &= \beta_1 + \beta_2 x_t + \varepsilon_t \\ &= \hat{\beta}_1 + \hat{\beta}_2 x_t + e_t \\ \Rightarrow & \quad e_t = y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t \end{aligned}$$

- Where, for all t

$$\begin{aligned} E(\epsilon_t) &= 0 \\ Var(\epsilon_t) &= \sigma^2 \end{aligned}$$

- y_t is drawn has a mean of $\beta_1 + \beta_2 x_t$ and a variance of σ^2.
- If these estimated errors are squared and summed we obtain

$$SSE = \sum_{t=1}^{n} e_t^2 = \sum_{t=1}^{n} (y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t)^2$$