W06

- Method
  - The Method of Moments
    - Let $X_1, \dots, X_n$ be a random sample from a pmf or pdf $f(x)$. For $k = 1, 2, 3, \dots$, the ***k*th population moment, or *k*th moment of the distribution** $f(x)$, is $E(X^k)$. The ***k*th sample moment** is $(1/n)\sum_{i=1}^{n} X_i^k$.
    - First
      - Population
        - $E(X) = \mu$.
      - Sample
        - $\Sigma X_i / n = \bar{X}$.
    - Let $X_1, X_2, \dots, X_n$ be a random sample from a distribution with pmf or pdf $f(x; \theta_1, \dots, \theta_m)$, where $\theta_1, \dots, \theta_m$ are parameters whose values are unknown. Then the **moment estimators** $\hat{\theta}_1, \dots, \hat{\theta}_m$ are obtained by equating the first m sample moments to the corresponding first m population moments and solving for $\theta_1, \dots, \theta_m$.
    - Second
    - exponential
      - $\hat{\lambda} = 1/\bar{X}$.
    - gamma distribution
      - $\bar{X} = \alpha\beta \qquad \frac{1}{n}\sum X_i^2 = \alpha(\alpha + 1)\beta^2$
      - $\hat{\alpha} = \frac{\bar{X}^2}{(1/n)\Sigma X_i^2 - \bar{X}^2} \qquad \hat{\beta} = \frac{(1/n)\Sigma X_i^2 - \bar{X}^2}{\bar{X}}$
    - negative binomial
      - $\hat{p} = \frac{\bar{X}}{(1/n)\Sigma X_i^2 - \bar{X}^2} \qquad \hat{r} = \frac{\bar{X}^2}{(1/n)\Sigma X_i^2 - \bar{X}^2 - \bar{X}}$
  - Maximum Likelihood Estimation
    - Natural log of the joint pmf is often easier
    - Let $X_1, X_2, \dots, X_n$ have joint pmf or pdf
      $$f(x_1, x_2, \dots, x_n; \theta_1, \dots, \theta_m) \tag{6.6}$$
      where the parameters $\theta_1, \dots, \theta_m$ have unknown values. When $x_1, \dots, x_n$ are the observed sample values and (6.6) is regarded as a function of $\theta_1, \dots, \theta_m$, it is called the **likelihood function.** The maximum likelihood estimates (mle's) $\hat{\theta}_1, \dots, \hat{\theta}_m$ are those values of the θ_i's that maximize the likelihood function, so that
      $$f(x_1, \dots, x_n; \hat{\theta}_1, \dots, \hat{\theta}_m) \ge f(x_1, \dots, x_n; \theta_1, \dots, \theta_m) \text{ for all } \theta_1, \dots, \theta_m$$
      When the X_i's are substituted in place of the x_i's, the **maximum likelihood estimators** result.
      - i.i.d
    - exponential
      - $\lambda = n/\sum x_i = 1/\bar{x}$
    - Normal Distribution
      - $\ln[f(x_1, \dots, x_n; \mu, \sigma^2)] = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum(x_i - \mu)^2$
      - $\hat{\mu} = \bar{X} \qquad \hat{\sigma}^2 = \frac{\sum(X_i - \bar{X})^2}{n}$
        - sigma is biased
    - Poisson
      - $\hat{\lambda} = \sum X_i / \sum a(R_i)$.
    - Weibull
  - Estimating functions of parameters
    - The Invariance Principle
      Let $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$ be the mle's of the parameters $\theta_1, \theta_2, \dots, \theta_m$. Then the mle of any function $h(\theta_1, \theta_2, \dots, \theta_m)$ of these parameters is the function $h(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m)$ of the mle's.
  - Large Sample Behavior of the MLE
    - Under very general conditions on the joint distribution of the sample, when the sample size n is large, the maximum likelihood estimator of any parameter θ is approximately unbiased $[E(\hat{\theta}) \approx \theta]$ and has variance that is either as small as or nearly as small as can be achieved by any estimator. Stated another way, the mle $\hat{\theta}$ is approximately the MVUE of θ.
  - Some Complications
- Excel
  - Add-on: Solver
- Concept
  - point estimate
    - $\hat{\theta}$ ("theta hat")
    - $\hat{\theta} = \theta + \text{error of estimation}$
    - *expected* or *mean square error* $\text{MSE} = E[(\hat{\theta} - \theta)^2]$.
    - Trade off: Bias? Efficient?
  - Unbiased estimator
    - **unbiased estimator** of θ if $E(\hat{\theta}) = \theta$
      - Center: Expected value, not mean or medium
      - Prefer unbiased estiminator
      - $\hat{\sigma}^2 = S^2 = \frac{\sum(X_i - \bar{X})^2}{n - 1}$
        - $V(Y) = E(Y^2) - [E(Y)]^2$,
      - $\tilde{X}$
      - $\bar{X}$ and any trimmed mean
        - continuous and symmetric
      - Bias: E[theta_hat] - theta
    - Binomial, n and p
      - $\hat{p} =$ X/n is an unbiased estimator of p.
      - $\sigma_{\hat{p}} = \sqrt{V(X/n)} = \sqrt{\frac{V(X)}{n^2}} = \sqrt{\frac{npq}{n^2}} = \sqrt{\frac{pq}{n}}$
    - Uniform distribution
      - the interval from 0 to an unknown upper limit θ
      - $\hat{\theta}_1 = \max(X_1, \dots, X_n)$.
      - $\hat{\theta}_2 = \frac{n+1}{n} \cdot \max(X_1, \dots, X_n)$
        - unbiased
      - $V(X) = \sigma^2 = (B - A)^2/12$.
    - Normal distribution
      - $\hat{\sigma}^2 = S^2 = \frac{\sum(X_i - \bar{X})^2}{n - 1}$
        - unbiased
      - $\hat{\sigma}^2 = \frac{\sum(X_i - \bar{X})^2}{n}$
        - Preferred, from MLE
      - $\hat{\mu} = \bar{X}$ is the MVUE for μ.
  - Estimator with min Variance
    - **MVUE: minimum variance unbiased estimator**
  - Complicated situation
    - The best estimator for mu depends on distribution
    - Exponential
      - parameter λ so that expected lifetime is $\mu = 1/\lambda$.
      - $T_r = \sum_{i=1}^{r} Y_i + (n - r)Y_r$
      - $\hat{\mu} = T_r/r$
      - $V(T_r/r) = 1/(\lambda^2 r)$,
  - The standard error
    - $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$.
    - estimated standard error
      - normal distribution: 2 sigma
      - 4 sigma
      - Boostrap