# Lecture 30:
# Estimation - Part III

Satyajit Thakor
IIT Mandi

# Recall: maximum likelihood estimation

- How to find maximum likelihood estimation of a parameter?
- Step 1: Write the expression for the likelihood function for the given distribution.
- Step 2: If the expression is complex, take logarithm both sides.
- Step 3: Differentiate the likelihood function (or log-likelihood function) with respect to the parameter to be estimated.
- Step 4: Equate it to zero to find the estimate(s) of the parameter which maximizes the likelihood function.
- Step 5: Conduct the second derivative test - if the second derivative at an estimate obtained in Step 4 is negative then the estimate provides a (local) maximum value of the function.

# Maximum likelihood estimation

- Example: Given a sample $x_1, \ldots, x_n$ find maximum likelihood estimator of $\lambda$ for $X \sim \text{Pois}(\lambda)$. (Recall Poisson distribution: Lecture 9)

- Recall: $X \sim \text{Pois}(\lambda)$ : $P(X=x) = \dfrac{e^{-\lambda}\lambda^{x}}{x!}$, $x = 0, 1, 2, \ldots$

- Then, likelihood function is

$$L(x_1, \ldots, x_n; \lambda) = \prod_{i=1}^{n} f(x_i; \lambda)$$

$$= \prod_{i=1}^{n} P(X_i = x_i)$$

$$= \frac{e^{-n\lambda} \lambda^{\sum x_i}}{\prod x_i!}$$

# Maximum likelihood estimation

- The log-likelihood function is

$$\log_e L(x_1,\dots,x_n;\lambda) = \log_e\left(\frac{e^{-n\lambda}\,\lambda^{\sum x_i}}{\prod x_i!}\right)$$

$$= -n\lambda + \sum x_i \log_e \lambda - \log_e\left(\prod x_i!\right).$$

- Differentiate:

$$\frac{d}{d\lambda}\log_e L(x_1,\dots,x_n;\lambda) = -n + \frac{\sum x_i}{\lambda} - 0 .$$

- Equate to zero: $-n + \frac{\sum x_i}{\hat{\lambda}} = 0 \Rightarrow \hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n}.$

# Maximum likelihood estimation

- Second derivative test:

$$\frac{d^2}{d\lambda^2}\log_e L(x_1,\dots,x_n;\lambda)\bigg|_{\lambda=\hat\lambda} = -\frac{\sum x_i}{\lambda^2}\bigg|_{\lambda=\frac{\sum_{i=1}^n x_i}{n}}$$

$$= -\frac{\sum x_i}{(\sum x_i)^2/n^2}$$

$$= -\frac{n^2}{\sum_{i=1}^n x_i} < 0.$$

$\Rightarrow \hat\lambda = \dfrac{\sum_{i=1}^n x_i}{n}$ maximizes the log-likelihood function and hence it is the maximum likelihood estimate.

# Interval estimation

- So far we studied point estimation of a parameter.
- For example, sample mean, sample variance, unbiased point estimators, maximum likelihood estimators.
- A point estimate $\bar{x}$ of the mean for a given sample $x_1, \ldots, x_n$ is not always the parameter $\mu$ it estimates but it is "very close" to $\mu$.

- Hence, rather than a point estimate, it is sometimes more valuable to be able to specify an interval for which we have a certain degree of confidence that $\mu$ lies within.
- To obtain such an interval estimator, we make use of the probability distribution of the point estimator $\bar{X}$.

# Interval estimation

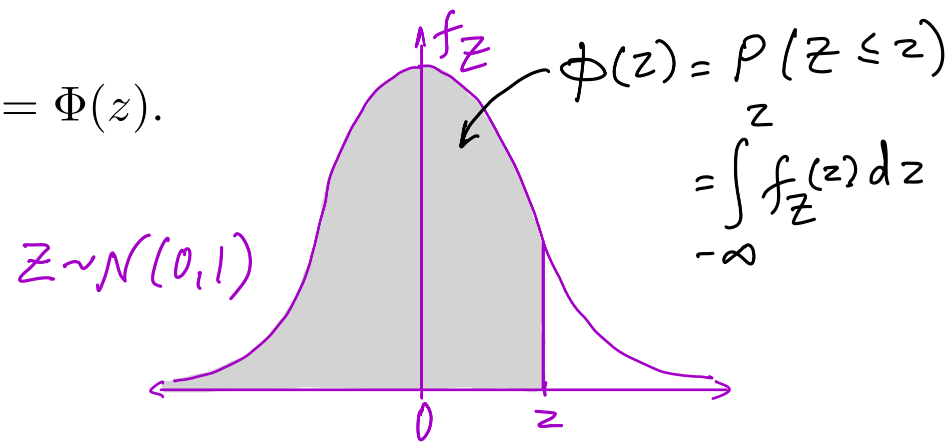
- An interval estimate of a population parameter $\theta$ is an interval of the form $\hat{\theta}_L < \theta < \hat{\theta}_U$, where $\hat{\theta}_L$ and $\hat{\theta}_U$ depend on the value (e.g., $\hat{\theta}$) of the statistic $\hat{\Theta}$ for a particular sample and also on the distribution of the parameter $\hat{\Theta}$.
- If, for instance, we find $\hat{\theta}_L$ and $\hat{\theta}_U$ such that

$$P(\underbrace{\hat{\theta}_L}_{\text{lower}} < \theta < \underbrace{\hat{\theta}_U}_{\text{upper}}) = 1 - \alpha,$$

  for $0 < \alpha < 1$, then we have a probability of $1 - \alpha$ of selecting a random sample that will produce an interval containing $\theta$.
- In this case, the interval $\hat{\theta}_L < \theta < \hat{\theta}_U$ is called $100(1 - \alpha)$ percent confidence interval estimate of $\theta$.
- Next we will try to find $100(1 - \alpha)$ percent confidence interval estimate of a sample mean $\bar{X}$.

# Interval estimation

- For an example of $100(1-\alpha)$ percent confidence interval estimate of a sample mean $\bar{X}$, assume that $\bar{X}$ is normally distributed.
- Note that this is a very logical assumption for large $n$ since the CLT suggests that the distribution of $\bar{X}$ can be well approximated by the normal distribution $\mathcal{N}(\mu, \sigma^2/n)$ (recall: Slides 4-5 of Lecture 26).
- Also, recall that if $\bar{X}$ has the distribution $\mathcal{N}(\mu, \sigma^2/n)$ then $\bar{X} = \mu + (\sigma/\sqrt{n})Z$, where, $Z \sim \mathcal{N}(0,1)$ (recall: Slide 7 of Lecture 25).
- Now, denote $P(Z \leq z) = \Phi(z)$.



# Interval estimation

- Let $z_\alpha$ be the value such that $\Phi(z_\alpha) = 1 - \alpha$.
- For any give $0 < \alpha < 1$, we can find out $z_\alpha$ numerically (using the CDF table for standard normal).