Central Limit Theorem

- $y_1, \ldots y_n$ are drawn from a distribution **independently** with finite mean μ and variance σ^2, then

$$\mathbb{E}\bar{y} = \mu, \quad \mathrm{Var}(\bar{y}) = \sigma^2/n.$$

- **Central Limit Theorem** When n is sufficiently large, $\bar{y}$ can be approximated by a normal distribution with mean μ and variance σ^2/n, i.e.,

$$\frac{\bar{y} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

- The sampling distribution of a sum of random variables

$$\frac{\sum_{i=1}^{n} y_i - n\mu}{n\sigma^2} \sim N(0, 1)$$

Normal Approximation to Binomial

- $y \sim \text{Bi}(n, p)$, i.e.

$$y = \sum_{i=1}^{n} y_i, \quad y_i = 0, 1 \sim \text{Bernoulli}(p)$$

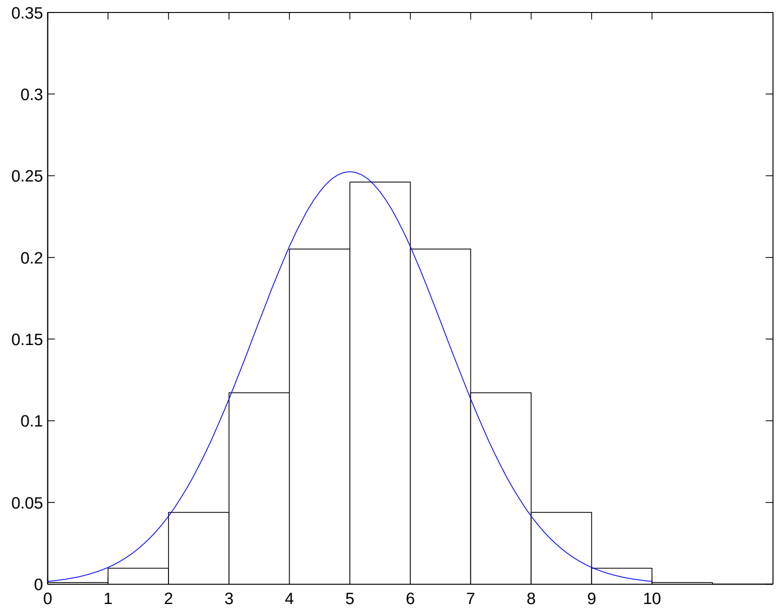
- $\mathbb{E}y = np, \quad \text{Var}(y) = npq$

- By Central Limiting Theorem,

$$\frac{y - np}{\sqrt{npq}} \sim N(0, 1)$$

$$P(y > np + x\sqrt{npq}) \approx \int_{\infty}^{x} \frac{\exp(-t^2/2)}{\sqrt{2\pi}} dt$$

- The approximation will be good if both $np \geq 4$ and $nq \geq 4$.



Continuity Correction for the Normal Approximation to a Binomial Probability

Let y be a Bin(n, p) and let $z = (y - np)/\sqrt{npq}$. Then

$$
\begin{aligned}
P(y \leq a) &= P(y < a + 0.5) \approx P(z < \frac{a\mathbf{+0.5} - np}{\sqrt{npq}}) \\
P(y \geq a) &= P(y > a - 0.5) \approx P(z > \frac{a\mathbf{-0.5} - np}{\sqrt{npq}}) \\
P(a \leq y \leq b) &= P(a - 0.5 < y < b + 0.5) \\
&\approx P(\frac{a\mathbf{-0.5} - np}{\sqrt{npq}} < z < \frac{b\mathbf{+0.5} - np}{\sqrt{npq}})
\end{aligned}
$$

Example: Height of Plants

Without an automated irrigation system, the height of plants two weeks after germination is normally distributed with a mean of 2.5 centimeters and a standard deviation of 0.5 centimeters.

It is reasonable to assume that with an automated irrigation system, the height of plants two weeks after germination is also normally distributed.

How to guess the mean? (*parameter estimation*)

Does the automated irrigation system have effects on the heights of plants? (*hypothesis testing*)

Concepts of Point Estimate

- A **point estimate** of some population parameter θ is a numerical value calculated from the sample.

- A **point estimator** is a formula or rule that tells us how to calculate a numerical estimate, denoted by $\widehat{\theta}(y_1, y_2, \ldots, y_n)$.

- The **bias B** of an estimator $\widehat{\theta}$ is equal to

$$B = \mathbb{E}(\widehat{\theta}) - \theta$$

- An estimator $\widehat{\theta}$ is **unbiased** if $\mathbb{E}(\widehat{\theta}) = \theta$, i.e., $B = 0$.

- The **mean squared error** of a point estimator is equal to

$$\begin{aligned} \mathbb{E}[(\hat{\theta}-\theta)^2] &= \mathbb{E}[(\hat{\theta}-\mathbb{E}(\hat{\theta})+\mathbb{E}(\hat{\theta})-\theta)^2] \\ &= \text{Var}(\hat{\theta})+B^2 \quad \text{why?} \end{aligned}$$

- The **minimum variance unbiased estimate (MVUE)** is the unbiased estimator $\hat{\theta}$ that has the smallest variance of all unbiased estimators.

Examples:

- (example 8.1 on page 341) Suppose $y_1, y_2, \ldots y_n$ be a random sample from some distribution with mean μ and variance σ^2. Show that the sample mean $\bar{y}$ and sample variance s^2 are unbiased estimators of μ and σ^2.

- (Exercise 8.3) Suppose y has a binomial distribution with parameter n and p. (1) Show that $\widehat{p} = y/n$ is an unbiased estimator of p; (2) Find the mean square error of the estimator $\widehat{p}$.

- (Exercise 8.6) Suppose $y \sim \text{Unif}(2, \theta)$. (1) Show that y_1 is a biased estimator of θ and compute the bias; (2) Show that $2(y_1 - 1)$ is an unbiased estimator of θ.

Method of Moment

- Let $y_1, y_2, \ldots, y_n$ represent a random sample of size n from some distribution.

  kth population moment: $\mathbb{E}(y^k)$

  kth sample moment:

  $$m^k = \frac{y_1^k + y_2^k + \cdots + y_n^k}{n}$$

- Supose the population distribution has parameters $\theta_1, \ldots, \theta_m$. Then the **moment estimators**, $\widehat{\theta}_1, \ldots, \widehat{\theta}_m$, are obtained by equating the first m sample moments to the corresponding first m population moments and solving the resulting equations for the unknown parameters.

Examples: Find the moment estimators.

- $y_1, y_2, \ldots, y_n \sim$ Exp(β). ($\widehat{\beta} = \bar{y}$)

- $y_1, y_2, \ldots, y_n \sim$ Poisson(λ). ($\widehat{\lambda} = \bar{y}$)

- $y_1, y_2, \ldots, y_n \sim$ Gamma(α, β). (see example 8.3 on page 346)

- $y_1, y_2, \ldots, y_n \sim N(\mu, \sigma^2)$.

$$\widehat{\mu} = \bar{y} \quad \widehat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n}.$$

Note: The moment estimator of σ^2 is not an unbiased estimator.

Method of Maximum Likelihood

- Suppose we randomly select a sample of n observations, $y_1, \ldots, y_n$ from a distribution $p(y \mid \theta)$, where θ is an unknown parameter. Then the **likelihood** of the sample is

$$L(\theta) = p(y_1 \mid \theta) \cdot p(y_2 \mid \theta) \cdots p(y_n \mid \theta)$$

  Note:

  - The likelihood is the joint probability function $p(y_1, \ldots, y_n \mid \theta)$ when y_i's are discrete r.v.

  - The likelihood is the joint density function $f(y_1, \ldots, y_n \mid \theta)$ when y_i's are continuous r.v.

  - Once we have observed y_i's, the likelihood function is a function of only the unknown parameter θ.

- **Ronald A. Fisher** (1890-1962):

  One should choose as an estimate of θ the value of θ that maximizes the likelihood $L(\theta)$.

- MLE $\widehat{\theta} = \text{argmax}_\theta \; L(\theta)$

- MLE $\widehat{\theta} = \text{argmax}_\theta \; \log L(\theta)$

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} p(y_i \mid \theta) \\ \log L(\theta) &= \sum_{i=1}^{n} \log p(y_i \mid \theta) \end{aligned}$$

How to Find MLE

- MLE $\hat{\theta} = \text{argmax}_\theta \; L(\theta)$.

- Solve
$$\frac{dL}{d\theta} = 0 \qquad \text{or} \quad \frac{d \log L}{d\theta} = 0$$
and check that the resulting solution is a maximum.

- Examples:

  (1) (example 8.4) Find the MLE of Exp(β).

  (2) (example 8.5) Find the MLE of Normal(μ, σ^2).

Complications in Using MLE

- Maximum occurs at a discontinuous point.

  Example: Suppose $y1, \ldots, y_n \sim$ Unif$(0, \theta)$. Find the MLE of θ.

- Close-form solution does not exist

  Example: Suppose $y1, \ldots, y_n \sim$ Gamma(α, β). Find the MLE.