STAT 135 Lecture 3

Henry Liev

26 June 2025

Example 0.1

Estimating population variance:
$Y = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$
$\frac{1}{n}\sum_{i=1}^{n}(x_i - c)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 + (\bar{x} - c)^2$ minimized for $c = \bar{x}$
$c = \mu$ $\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 + (\bar{x} - \mu)^2$
$\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 - (\bar{x} - \mu)^2$
$E(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2) = E(\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 - (\bar{x} - \mu)^2) = \sigma^2 - \frac{\sigma^2}{n} = \frac{(n-1)\sigma^2}{n}$
To remove the bias, we need to multiply by $\frac{n}{n-1}$:
$\sigma^2 = E(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)$

Example 0.2

Fitting Probability Distributions to Data:
$X_1, X_2, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$
$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Example 0.3

Degrees of Assumptions:

1. Finite population
   Flipping a biased coin, probability of heads, can't be wrong
2. Very good reason to believe the model
   number of calls coming to a call center in one hour, Poisson: $P(x = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
3. Could be true, Gamma distribution mode, no strong theory

Definition 0.4

Moments:
k^{th} moments: $\mu_k = E(X^k)$
$\mu_1 = E(X)$ 'location' $\hat{\mu_1} = \frac{1}{n}\sum x_i$
$\mu_2 = E(X^2)$ 'spread' $\hat{\mu_2} = \frac{1}{n}\sum x_i^2$
$\mu_3 = E(X^3)$ 'skewness/symmetry' $\hat{\mu_3} = \frac{1}{n}\sum x_i^3$
$\mu_4 = E(X^4)$ 'tail weights' $\hat{\mu_4} = \frac{1}{n}\sum x_i^4$

Method 0.5

Method of Moments

1. Calculate some moments of the distribution (formulas in terms of parameters, compute as many as number of parameters)
2. Solve for parameters in terms of moments
3. Plug in $\hat{\mu_k}$ for μ_k

Example 0.6

$x_1, x_2, \ldots, x_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$
$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
$\mu_1 = E(X) = \mu$
$\mu_2 = E(X^2) = \mu^2 + \sigma^2 \to \sigma^2 = \mu_2 - \mu_1^2$
$\hat{\mu}_1 = \bar{x}$ Estimator of μ
$\mu_2 - \mu_1^2 = \frac{1}{n} \sum x_i^2 - (\bar{x})^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$ Estimator of σ^2

Remark 0.7

Consistency of MoM Estimators:

1. MoM are consistent (converges to the right value in probability)
2. $\bar{x} \xrightarrow{p} \mu$ (Weak Law of Large Numbers)
   $\forall \varepsilon > 0, P(|\bar{x} - \mu| > \varepsilon) \to 0$ as $n \to \infty$
   True because of Chebyshev's inequality:
   $P(|\bar{x} - \mu| > \varepsilon) \leq \frac{Var(\bar{x})}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$
3. Also true that $\hat{\mu}_k \xrightarrow{p} \mu_k, \forall k$
   $f(\hat{\mu}_k) \xrightarrow{p} f(\mu_k)$ for continuous f
4. Parameter estimates are usually continuous functions of the moments (Continuous Approximation Theorem)

Example 0.8

Let $x_1, x_2, \ldots, x_n$ from a discrete probability distribution $P(X = 0) = \frac{2}{3}\theta, P(X = 1) = \frac{1}{3}\theta, P(X = 2) = \frac{2}{3}(1-\theta), P(X = 3) = \frac{1}{3}(1-\theta)$
$\mu_1 = 0 \cdot \frac{2}{3}\theta + 1 \cdot \frac{1}{3}\theta + 2 \cdot \frac{2}{3}(1-\theta) + 3 \cdot \frac{1}{3}(1-\theta) = \frac{7}{3} - 2\theta$
$\theta = \frac{1}{2}(\frac{7}{3} - \mu_1)$
$\hat{\theta} = \frac{1}{2}(\frac{7}{3} - \hat{\mu}) = \frac{1}{2}(\frac{7}{3} - \bar{x})$