STAT 135 Lecture 7

Henry Liev

8 July 2025

Remark 0.1

Objectives: Sufficiency, Rao-Blackwell Theorem
Estimation:

- Construction: MOM, MLE
- Evaluation: Consistency, Unbiased
- Optimality

Example 0.2

Coin Flipping
Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$
Estimate p
Intuitively, number of heads is all you need to estimate p
Intuitevely, $\sum X_i$ contains all the information in the data about p
Conceptually, a sufficient statistic for θ is one which contains all the information

Remark 0.3

A **Sufficient Statistic** for θ is one which contains all the information

- It is more efficient for storage
- For privacy reasons, aggregates could be better
- Guidance for forming estimators

Definition 0.4 (Sufficiency)

Data $X_1, \ldots, X_n \sim f(x|\theta)$
A statistic T is **sufficient** for θ if the conditional distribution of $X_1, \ldots, X_n$ given $T = t$ does not depend on θ for any t

Remark 0.5

Why does this make sense?

- After conditioning on T conditional distributions of X's does not depend on θ.
- No further analysis after conditioning on T can tell you about θ.

Example 0.6

$X_1, X_2, \dots, X_n \overset{iid}{\sim} Bernoulli(p)$

Solution. $\sum_{i=1}^n X_i$ is a sufficient statistic for p.
Need to show: $P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | T = t) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | \sum_{i=1}^n X_i = t)$ doesn't depend on p

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | \sum_{i=1}^n X_i = t) = \frac{P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n \cap \sum_{i=1}^n x_i = t)}{P(\sum_{i=1}^n x_i = t)}$$
$$= \frac{p^t(1-p)^{n-t}}{\binom{n}{t} p^t (1-p)^{n-t}} = \frac{1}{\binom{n}{t}}$$

■

Theorem 0.7 (Factorization Criterion)

Easy way to check/identify sufficient statistics
$X_1, \dots, X_n \sim f(x|\theta)$
A statistic $T(X_1, \dots, X_n)$ is sufficient for $\theta \Leftrightarrow$

$$f(x_1, \dots, x_n) = g(T(x_1, \dots, x_n), \theta) \cdot h(x_1, \dots, x_n)$$

Method 0.8 (Playbook for factorization criterion)

1. Write down joint density
2. Check if it factors this way

Example 0.9

$X_1, \dots, X_n \overset{iid}{\sim} Poisson(\lambda)$
Show $T = \sum_{i=1}^n X_i$ is sufficient for λ

Solution.

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n \frac{1}{x_i!} \lambda^{x_i} e^{-\lambda} = \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \prod_{i=1}^n \frac{1}{x_i!}$$
$$g(T, \lambda) \lambda^T e^{-n\lambda}$$
$$h(x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{x_i!}$$

■

Remark 0.10

More reasons to care about sufficient statistics

1. $\hat{\theta}_{MLE}$ is always a function of a sufficient statistics
   Suppose we have data $X_1, \ldots, X_n \sim f(X_1, \ldots, X_n|\theta)$ and T is sufficient for θ then by the factorization criterion $f(x_1, \ldots, x_n|\theta) = g(T, \theta) \cdot h(x_1, \ldots, x_n)$

$$\hat{\theta}_{MLE} = \arg\max_{\theta} f(x_1, \ldots, x_n|\theta) = \arg\max_{\theta} g(T, \theta)$$

Theorem 0.11 (Rao-Blackwell Theorem)

$x_1, \ldots x_n \sim f(x_1, \ldots, x_n|\theta)$
Have an estimator $\hat{\theta}$ of θ (Need to assume $E(\hat{\theta}^2) < \infty$, finite second moment)
Assume we have a sufficient statistic T for θ
Create $\tilde{\theta} = E\left[\hat{\theta}|T\right]$ (Random Variable)
$MSE(\tilde{\theta}) \leq MSE(\hat{\theta})$

Proof. $\hat{\theta}$ original
$\tilde{\theta} = E\left[\hat{\theta}|T\right]$ new and improved
$E\left[E(\hat{\theta}|T)\right] = E(\hat{\theta})$
$E(\tilde{\theta}) = E(\hat{\theta})$ Have the same bias
Focus on variance $Var(\hat{\theta}) = Var\left[E(\hat{\theta}|T)\right] + E\left[Var(\hat{\theta}|T)\right] = Var(\tilde{\theta}) + E\left[Var(\hat{\theta}|T)\right]$
$E\left[Var(\hat{\theta}|T)\right] \geq 0$ ■