Lecture Summary: Errors in Parameter Estimation

Lecture: 9.3 - Errors in Parameter Estimation

Source: Lec8.3.pdf

Key Points

- **The Parameter Estimation Problem:**
  - Estimation involves deriving an unknown parameter θ from iid samples $X_1, X_2, \ldots, X_n$.
  - The parameter θ is fixed, but the estimator $\hat{\theta}$ is a random variable with its own distribution.
  - A good estimator produces errors that are small and close to zero.
- **Error in Estimation:**
  - Error is defined as:
  $$\text{Error} = \hat{\theta}(X_1, \ldots, X_n) - \theta.$$
  - The error is a random variable, and its absolute value should ideally remain small.
  - A mathematical approach to controlling error:
  $$P(|\text{Error}| > \delta) \text{ should be small, where } \delta \text{ is a threshold.}$$
  - The choice of δ is context-dependent and often relative to the magnitude of θ.
- **Relative Error Thresholds:**
  - Errors should be characterized as a fraction of the parameter being estimated.
  - For example:
    * For Bernoulli(p): $|\text{Error}| \leq p/10$ (10% error relative to p).
    * For Normal(μ, σ^2): Error relative to μ varies with scale.
- **Comparing Estimators:**
  - Three estimators for p in Bernoulli trials were evaluated:
    1. $\hat{p}_1 = 0.5$ (fixed).
    2. $\hat{p}_2 = \frac{X_1+X_2}{2}$ (uses only the first two samples).
    3. $\hat{p}_3 = \frac{\sum_{i=1}^{n} X_i}{n}$ (sample mean).
  - Observations:
    * $\hat{p}_1$ is constant and does not adapt to p.
    * $\hat{p}_2$ adapts but has high variability.
    * $\hat{p}_3$ uses all samples, balances adaptability and stability, and shows the best performance.
- **Chebyshev's Inequality in Estimation:**

- Chebyshev's bound on error:

$$P(|\text{Error}| > \delta) \leq \frac{\text{Var}(\text{Error})}{\delta^2}.$$

- For $\hat{p}_3$:
  * $\mathbb{E}[\text{Error}] = 0$ (unbiased).
  * $\text{Var}(\text{Error}) = \frac{p(1-p)}{n}$.
  * Probability bound for $|\text{Error}| > p/10$:

$$P(|\text{Error}| > p/10) \leq \frac{100(1-p)}{np}.$$

- As $n \to \infty$, $P(|\text{Error}| > \delta) \to 0$, demonstrating the estimator's performance improvement with more samples.

- **Key Insights:**
  - Good estimators adapt to the parameter and leverage all available data.
  - Increasing the sample size reduces error variance and improves reliability.
  - Concentration results like Chebyshev's and Chernoff bounds highlight how probabilities of large errors diminish with more samples.

Simplified Explanation

Key Idea: Errors in parameter estimation should be small and decrease as the sample size increases.

Comparison of Estimators: - $\hat{p}_3$ (sample mean) is effective because it uses all samples, adapts to p, and reduces error variance as n grows.

Why It Matters: Accurate estimators provide reliable parameter estimates for decision-making and data analysis.

Conclusion

In this lecture, we:

- Explored how errors in parameter estimation are characterized and controlled.
- Evaluated the performance of different estimators for Bernoulli(p).
- Used Chebyshev's inequality to quantify error probabilities.

The concepts discussed are foundational for designing effective estimators that leverage data efficiently while minimizing estimation errors.