Discussion 3

- **Prediction**
  - *(Best MSPE predictor)* $\mu(Z) = E(Y|Z)$
  - *(Best linear predictor)* $\mu(Z) = a_1 + b_1 Z$ where

$$b_1 = \frac{\text{Cov}(Z,Y)}{\text{Var}(Z)}, \quad a_1 = E(Y) - b_1 E(Z)$$

  - *(Best zero intercept linear predictor)* $\mu(Z) = b_0 Z$ where

$$b_0 = \frac{E(ZY)}{E(Z^2)}$$

- **Sufficient Statistics**
  - A statistics $T(X)$ is a *sufficient* for $P \in \mathcal{P}$ or the parameter θ if the conditional distribution of the sample X given the value of $T(X) = t$ does not depend on θ.
  - **(Factorization Theorem)** In a regular model, a statistic $T(X)$ is sufficient for θ if and only if there exist functions $g(t, \theta)$ and $h(x)$ such that for all $x \in \mathcal{X}$, $\theta \in \Theta$

$$p(x, \theta) = g(T(x), \theta) h(x)$$

1.4.1

An urn contains 4 red and 4 black balls. 4 balls are drawn at random without replacement. Let Z be the number of red balls obtained in the first 2 draws and Y the total number of red balls drawn.

1. Find the best predictor of Y given Z, the best linear predictor, and the best zero intercept linear predictor.
2. Compute the MSPEs of the predictors in (a).

1.5.1

Let $X_1, \cdots, X_n$ be a sample from a Poisson, $\mathcal{P}(\theta)$, population where $\theta > 0$.

1. Show directly that $\sum_{i=1}^n X_i$ is sufficient for θ.
2. Establish the same result using the factorization theorem.

1.5.3

Suppose $X_1, \cdots, X_n$ is a sample from a population with one of the following densities.

1. $p(x, \theta) = \theta x^{\theta - 1}, 0 < x < 1, \theta > 0$
2. $p(x, \theta) = \theta a x^{a-1} \exp(-\theta x^a), x > 0, \theta > 0, a > 0$

In each case, find a real-valued sufficient statistic for θ, a fixed.