Summary: We are considering taking possible actions $a \in \mathcal{A}$, and unknown quantities affecting our decision are represented by $\theta \in \Theta$.

In the estimation context, the action is just an estimate of θ, $\hat{\theta}(x)$.

The loss function describes the consequences of taking action a when the true state of nature is θ. We write it $L(\theta, a)$ or $L(\theta, \hat{\theta}(x))$.

Ultimately, we want to *choose* an action a or an estimate $\hat{\theta}(x)$. Our choice is driven by looking at a particular *risk function.*

So far we have seen two strategy:
(1) the Bayes rule, which chooses $\hat{\theta}(x)$ to minimize the Bayes risk; (2) the minimax rule, which minimizes the maximum frequentist risk.

Recall that for an estimation problem in a parametric framework, we also learned MOM and MLE.

Example: Suppose $X|p \sim Bin(n, p)$ and the loss is squared error.

- Show $\hat{p} = X/n$ is not minimax. *Hint: Consider the randomized estimator*

$$\tilde{p} = \left\{ \begin{array}{ll} X/n & \text{with probability } 1 - \frac{1}{n+1} \\ 1/2 & \text{with probability } \frac{1}{n+1} \end{array} \right\}$$

- Consider the Bayes estimator when $p \sim Beta(a, b)$. Find a and b so that the Bayes estimator has constant frequentist risk. This estimator is then minimax.

Example: Consider a decision problem with possible states of nature θ_1 and θ_2 . Let X be a random variable with probability function $p(x|\theta)$:
$P(X = 0|\theta_1) = 0.2, P(X = 1|\theta_1) = 0.8;$
$P(X = 0|\theta_2) = 0.4, P(X = 1|\theta_2) = 0.6.$

Two non-randomized actions a_1 and a_2 are considered with the following loss function:
$L(\theta_1, a_1(0)) = 1, L(\theta_1, a_1(1)) = 2, L(\theta_1, a_2(0)) = 4, L(\theta_1, a_2(1)) = 0;$
$L(\theta_2, a_1(0)) = 3, L(\theta_2, a_1(1)) = 1, L(\theta_2, a_2(0)) = 1, L(\theta_2, a_2(1)) = 4.$

1. Give and plot the risk set $S = \{(r_1, r_2) : r_1 = \lambda R(\theta_1, a_1) + (1 - \lambda)R(\theta_1, a_2), r_2 = \lambda R(\theta_2, a_1) + (1 - \lambda)R(\theta_2, a_2), \lambda \in [0, 1]\}$.

2. Suppose θ has the prior distribution $\Lambda(\theta)$ defined by $P(\theta = \theta_1) = 0.9, P(\theta = \theta_2) = 0.1$. What is the Bayes rule with respect to $\Lambda(\theta)$?

3. Find the minimax rule(s).

Review Question 1. Suppose we take a random sample of size n from a population of people. Let X_1 denote the number of individuals with a particular genotype AA, X_2 denote the number with Aa, and X_3 denote the number with aa. Assuming the gene frequencies are in equilibrium, the Hardy-Weinberg law says that the genotypes AA, Aa, and aa occur with probability $p_1 = (1-\theta)^2$, $p_2 = 2\theta(1-\theta)$, and $p_3 = \theta^2$, respectively.

1. What is the likelihood function for θ, treating $X = (X_1, X_2, X_3)$ as a sample from the multinomial distribution with size n and $p = (p_1, p_2, p_3)$?

2. Find the maximum likelihood estimator (MLE) for θ under this model. Find the aymptotic distribution (after appropriate normalization) for the MLE.

3. Construct the likelihood ratio test statistic for the null hypothesis that $p_1 = (1-\theta)^2$, $p_2 = 2\theta(1-\theta)$, and $p_3 = \theta^2$. What is the asymptotic distribution of your test statistic under null?

Review Question 2. Assume that we want to integrate $r(x)$ using a Monte Carlo Integration approach. How to select the density distribution $g(x)$ to sample $X_1, ..., X_B$? Can you assess the variance of your Monte Carlo integration result?

Remember that in Monte Carlo integration, we are finding the average of r/g for a number of sampled points. Hence if g is small for a given sample point, r/g will be arbitrarily large and can skew the sample mean from the true mean unless we generate a LARGE sample. One way to avoid such cases is to use g that looks like r. The peaks and valleys of g should correspond to peaks and valleys to r.

Review Question 3. Suppose $X_1, ..., X_n$ are iid Poisson(λ). Consider $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$.

1. Is S^2 an unbiased estimator for λ? Justify your answer.

2. Does S^2 have the smallest variance among all unbiased estimators for λ? If not, please find the unbiased estimator (for λ) that has the smallest variance. Justify your answer.

Review Question 4. Let be $(X_1, Y_1), ..., (X_n, Y_n)$ be iid with $X_i \sim N(0,1)$ and $Y_i|(X_i = x) \sim N(x\theta, 1)$.

1. Is the above an exponential family?

2. Find the fisher information $I(\theta)$.

3. Find an unbiased estimator of θ.