Discussion 2

- **Bayesian Model**
  - the prior $\pi(\theta)$ for θ.
  - the conditional distribution of X given θ, $p(x|\theta)$ (usually specified by model).
  - the posterior distribution of θ is simply the conditional distribution of θ given X

$$\pi(\theta|x) = \frac{f(x,\theta)}{p(x)} = \frac{\pi(\theta)p(x|\theta)}{p(x)} \propto \pi(\theta)p(x|\theta)$$

    where $p(x)$ is the marginal distribution of X.

- **Decision Theory**
  - **Action space** is the collection of all the actions or decisions, denoted by $\mathcal{A}$.
  - **Loss function** $l(P, a)$ is a function $l : \mathcal{P} \times \mathcal{A} \to \mathbb{R}^+$, which can be explained as the nonnegative loss incurred by taking action a, while the true distribution is P.
  - **Decision rule** δ is a function from the sample space $\mathcal{X}$ to the action space $\mathcal{A}$.
  - **Risk function** is a measure of the performance of the decision rule $\delta(X)$ by averaging loss over the sample space.

$$R(P, \delta) = E_P[l(P, \delta(X))]$$

    Particularly, if taking the quadratic loss $l(P, a) = (\nu(P) - a)^2$, then the risk function is called the *Mean Squared Error* (MSE) of $\hat{\nu}$.
  - **Bayes risk and Bayes rule:** The *Bayes risk* is defined as

$$r(\delta) = E_\Pi[R(\theta, \delta)] = E_{\Pi,P}[l(\theta, \delta(X))]$$

    and the *Bayes rule* is the decision rule that minimizes the Bayes risk. $r(\delta^*) = \min_\delta r(\delta)$
  - **Minimax rule** is the decision rule that minimizes the worst possible risk.

$$\sup_\theta R(\theta, \delta^*) = \inf_\delta \sup_\theta R(\theta, \delta)$$

1.2.1

Consider a parameter space consisting of two points θ_1 and θ_2, and suppose that for given θ, an experiment leads to a r.v. X whose frequency function $p(x|\theta)$ is given by

$\theta \setminus x$	0	1
θ_1	0.8	0.2
θ_2	0.4	0.6

Let π be the prior frequency function of θ defined by $\pi(\theta_1) = \pi(\theta_2) = \frac{1}{2}$.

(a) Find the posterior frequency function $\pi(\theta|x)$.

(b) Suppose $X_1, \cdots, X_n)$ are independent with frequency function $p(x|\theta)$. Find the posterior $\pi(\theta|x_1, \cdots, x_n)$. Observe that it depends only on $\sum_{i=1}^n x_i$.

(d) Give the values of $P(\theta = \theta_1 | \sum_{i=1}^n X_i = .5n)$ when n=2 and 100.

(e) Give the most probable value $\hat{\theta} = \arg\max_\theta \pi(\theta | \sum_{i=1}^n X_i = k)$ for n=2 and 100.

1.3.2

In example 1.3.5, the frequency function $p(x, \theta)$ is given by

	$x = 0$	$x = 1$
θ_1	0.3	0.7
θ_2	0.6	0.4

and all possible decision rules are

	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
$x = 0$	a_1	a_1	a_1	a_2	a_2	a_2	a_3	a_3	a_3
$x = 1$	a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3

Suppose a new buyer makes a bid and the loss function is

$\theta \setminus a$	a_1	a_2	a_3
θ_1	0	7	4
θ_2	12	1	6

(a) Compute and plot the risk points in this case for all possible nonrandomized rules.

(b) Find the minimax rule among the nonrandomized rules.

(c) Find the minimax rule among the randomized rules

(d) Suppose θ has prior $\pi(\theta_1) = \gamma$, $\pi(\theta_2) = 1 - \gamma$. Find the Bayes rule when (i) $\gamma = 0.5$.

1.3.9

Let θ denote the proportion of people working in a company who have a certain characteristic (e.g. being left-handed). It is known that in the state where the company is located, 10% have the characteristic. A person in charge of ordering equipment needs to estimate θ and uses

$$\hat{\theta} = (.2)(.10) + (.8)\hat{p}$$

where $\hat{p} = X/n$ is the proportion with the characteristic in a sample of size n from the company. Find $MSE(\hat{\theta})$ and $MSE(\hat{p})$. If the true θ is θ_0, for what θ_0 is

$$\frac{MSE(\hat{\theta})}{MSE(\hat{p})} < 1?$$

Give the answer for $n = 25$ and $n = 100$.