Problem 1: Bayesian Decision Rule
Solution:

1. a. $p(x)$ can be calculated based on the prior and the likelihood distributions as:

$$p(x) = p(x|\omega = 1)p(\omega = 1) + p(x|\omega = 2)p(\omega = 2)$$

for $x < 0$: $p(x) = 0$
for $0 \leq x < 1$: $p(x) = 0.5 \times 0.7 = 0.35$
for $1 \leq x \leq 2$: $p(x) = 0.5 \times 0.7 + 0.25 \times 0.3 = 0.425$
for $2 < x \leq 5$: $p(x) = 0.25 \times 0.3 = 0.075$
for $x > 5$: $p(x) = 0$

Therefore:

$$p(x) = \begin{cases} 0.35 & 0 \leq x < 1 \\ 0.425 & 1 \leq x \leq 2 \\ 0.075 & 2 < x \leq 5 \\ 0 & else \end{cases}$$

b. (1) The overall risk of Bayesian Decision Rule is:

$$\begin{aligned} R_{Bayes} &= \oint R(\alpha_{Bayes}(x)|x)p(x)dx \\ &= \oint \{1 - max[P(\omega = 1|x), P(\omega = 2|x)]\}p(x)dx \end{aligned}$$

Where

$$p(\omega = 1|x) = \frac{p(x|\omega = 1)p(\omega = 1)}{p(x)}$$

$$p(\omega = 2|x) = \frac{p(x|\omega = 2)p(\omega = 2)}{p(x)}$$

and

$$p(x|\omega = 1)p(\omega = 1) = \begin{cases} 0.35 & 0 \leq x \leq 2 \\ 0 & else \end{cases}$$

$$p(x|\omega = 2)p(\omega = 2) = \begin{cases} 0.075 & 1 \leq x \leq 5 \\ 0 & else \end{cases}$$

Thus,
for $0 \leq x < 1$: $p(\omega = 1|x) = 0.5 \times 0.7/p(x) = 1, p(\omega = 2|x) = 0$
for $1 \leq x \leq 2$: $p(\omega = 1|x) = 0.5 \times 0.7/p(x) = 35/42.5, p(\omega = 2|x) = 7.5/42.5$

for $2 < x \leq 5$: $p(\omega = 1|x) = 0, p(\omega = 2|x) = 1$

Therefore, according to $f(x)$ calculated in the previous problem, we can have the overall risk as:

$$\begin{aligned} R_{Bayes} &= \int_0^1 0 \times p(x)dx + \int_1^2 \{1 - p(\omega = 1|x)\}p(x)dx + \int_2^5 0 \times p(x)dx \\ &= \int_1^2 (7.5/42.5 \times 0.425)dx \\ &= 0.075 \end{aligned}$$

(2) Similarly, the overall risk for the Maximize Prior based Decision Rule is:

$$R_p = \oint R(\alpha_p(x)|x)p(x)dx$$

Where

$$\alpha_p(x) = \arg\max_{\omega} p(\omega) = 1$$

Then,

$$\begin{aligned} R_p &= \oint R(\omega = 1|x)p(x)dx \\ &= \oint p(\omega = 2|x)p(x)dx \\ &= \oint p(x|\omega = 2) \times p(\omega = 2)dx \\ &= \int_1^5 0.25 \times 0.3dx \\ &= 0.3 \end{aligned}$$

(3) The overall risk for the Maximize Likelihood based Decision Rule is:

$$R_L = \oint R(\alpha_L(x)|x)p(x)dx$$

Where

$$\alpha_L(x) = \arg\max_{\omega} p(x|\omega) = \begin{cases} 1 & 0 \leq x \leq 2 \\ 2 & 2 < x \leq 5 \end{cases}$$

Thus,

$$\begin{aligned} R_L &= \int_0^2 p(\omega = 2|x)p(x)dx + \int_2^5 p(\omega = 1|x)p(x)dx \\ &= \int_1^2 0.25 \times 0.3dx \\ &= 0.075 \end{aligned}$$

c. Both Bayesian Decision Rule and the Maximize Likelihood Decision Rule achieve the smallest overall risk, while the Maximize Prior Decision Rule has the largest overall risk.

2. The overall risk for Bayesian Decision Rule can be calculated as:

$$\begin{aligned} R_{Bayes} &= \oint R(\alpha_{Bayes}(x)|x)p(x)dx \\ &= \oint \{1 - max[P(\omega = j|x)|j = 1, ..., K]\}p(x)dx \end{aligned}$$

For simplicity, the class with max posterior probability is noted as $\omega_m ax$ and we can rewrite the overall risk as:

$$R_{Bayes} = \oint (1 - P(\omega_{max}|x))p(x)dx$$

For any decision rule, the overall risk can be calculated as:

$$\begin{aligned} R^* &= \oint R(\alpha(x)|x)p(x)dx \\ &= \oint \{\sum_j [P(\omega^*(x) \neq \omega_j|x)P(\omega_j|x)]\}p(x)dx \\ &= \oint \{\sum_j [(1 - P(\omega^*(x) = \omega_j|x))P(\omega_j|x)]\}p(x)dx \\ &= \oint \{1 - \sum_j [P(\omega^*(x) = \omega_j|x)P(\omega_j|x)]\}p(x)dx \end{aligned}$$

where $\omega^*(x) = \alpha(x)$.

Proving $R_{Bayes} \geq R^*$ under any distribution of x is equivalent to proving $\sum_j [P(\omega^*(x) = \omega_j|x)P(\omega_j|x)] \leq P(\omega_{max}|x)$:

$$\sum_j [P(\omega^*(x) = \omega_j|x)P(\omega_j|x)] \leq \sum_j [P(\omega^*(x) = \omega_j|x)P(\omega_{max}|x)] = P(\omega_{max}|x)$$

Thus proved that, Bayesian Decision Rule can always achieve the minimum overall risk.

Problem 2: Bayesian Reasoning

1. We would like to choose the cup which has the largest probability of having coffee in it. After the greeter opens cup C, there are only two cups A and B left. We can calculate the probability of having coffee in each cup and choose the larger one. The problem can be formulated as: Random variables $X = D_A, D_B, D_C$

Input data x is the cup that the greeter opened which has cold water in it. The prior probabilities are: $P(D_A), P(D_B), P(D_C)$ denoting the probabilities that A, B, C have a coffee in it, respectively.
The posterior probabilities are: $P(D_A|x), P(D_B|x), P(D_C|x)$ denoting the probabilities that A, B, C have a coffee in it after the greeter open a cup, respectively.

2. The prior is $P(D_A) = P(D_B) = P(D_C) = \frac{1}{3}$.

3. If A has coffee, the greeter can either open B or C, the likelihoods are:

$$p(x|D_A) = \begin{cases} \frac{1}{2} & x = B \\ \frac{1}{2} & x = C \end{cases}$$

If B or C has coffee, the greeter only has one option and the likelihood are:

$$p(x|D_B) = \begin{cases} 0 & x = B \\ 1 & x = C \end{cases}$$

$$p(x|D_C) = \begin{cases} 1 & x = B \\ 0 & x = C \end{cases}$$

4. The posterior probabilities given the input $x = C$ are:

$$\begin{aligned}
P(D_A|C) &= \frac{P(C|D_A)P(D_A)}{P(C)} \\
&= \frac{P(C|D_A)P(D_A)}{P(C|D_A)P(D_A) + P(C|D_B)P(D_B) + P(C|D_C)P(D_C)} \\
&= \frac{1/2 \cdot 1/3}{1/2 \cdot 1/3 + 1 \cdot 1/3 + 0 \cdot 1/3} \\
&= \frac{1}{3}. \\
P(D_B|C) &= \frac{P(C|D_B)P(D_B)}{P(C)} \\
&= \frac{P(C|D_A)P(D_A)}{P(C|D_A)P(D_A) + P(C|D_B)P(D_B) + P(C|D_C)P(D_C)} \\
&= \frac{1 \cdot 1/3}{1/2 \cdot 1/3 + 1 \cdot 1/3 + 0 \cdot 1/3} \\
&= \frac{2}{3}. \\
P(D_C|C) &= 0
\end{aligned}$$

5. The probability that A has hot coffee is $\frac{1}{3}$.

6. Yes, because the posterior probability that B has hot coffee is $\frac{2}{3}$, which is larger than the posterior probability that A has hot coffee.