Question Bank 3

Question 1

In a game show, you are asked to choose one of the three doors. One conceals a new car and two conceal goats. You choose, but your chosen door is not opened immediately. Instead the presenter opens another door, which reveals a goat. He approaches you and asks "would you like an opportunity to change your choice to the third door, which is unopened and unchosen so far?". Let p be the conditional probability that the third door conceals the car. The presenter's protocol is:

(i) he is determined to show you a goat, with a choice of two, he picks one at random. Show $p = \frac{2}{3}$.

(ii) he is determined to show you a goat; with a choice of two goats (named B and N) he shows you B with probability b. Show that, given you see b, the probability $p = 1/(1+b)$.

(iii) he opens a door chosen at random irrespective of what lies behind. Show $p = \frac{1}{2}$.

(iv) Show that, for $\alpha \in [\frac{1}{2}, \frac{2}{3}]$, there exists a protocol such that $p = \alpha$. Are you well advised to change your choice to the third door?

Solution: Given we cannot have two identical goats (doppelganger), we have in total 6 orderings of the goats and the car. Assume that each of the six orderings of the car and goats are equally likely. Let C_i be the event that the i^{th} door conceals the car, G the event that you see a goat, B the event that you see goat B in specific.

(i) According to the question, you are deemed to see a goat anyway. So the event we look at is $C_3|G$. Note that C_1 and C_3 are mutually exclusive events, therefore we will consider two scenarios: C_1 and C_1^c, i.e. first door conceals the car or not respectively.

$$P(C_3|G) = \frac{P(C_3 \cap G)}{P(G)} \tag{0.0.1}$$

$$= \frac{P(C_3 \cap G|C_1)P(C_1) + P(C_3 \cap G|C_1^c)P(C_1^c)}{P(G|C_1)P(C_1) + P(G|C_1^c)P(C_1^c)} \tag{0.0.2}$$

$$= \frac{0 \times \frac{1}{3} + 1 \times \frac{2}{3}}{1 \times \frac{1}{3} + 1 \times \frac{2}{3}} \tag{0.0.3}$$

Note: given the presenter determines to show a goat, $P(G|C_1) = P(G|C_1^c) = 1$.

(ii) Apply the similar idea with (i).

$$P(C_3|B) = \frac{P(C_3 \cap B)}{P(B)} \tag{0.0.4}$$

$$= \frac{P(C_3 \cap B|C_1)P(C_1) + P(C_3 \cap B|C_1^c)P(C_1^c)}{P(B|C_1)P(C_1) + P(B|C_1^c)P(C_1^c)} \tag{0.0.5}$$

Given C_1^c, that is no car behind the first door and there is a goat behind it, could be goat B or N with probability $\frac{1}{2}$ since we had the assumption that all these 6 orderings of the goats and the car are equally likely. Therefore $P(B|C_1^c) = \frac{1}{2}$. In other word, the second door has to be a goat, can be B or N with equal chance fifty-fifty. Now given C_1^c, it is certain for C_3 occurring. Therefore $P(C_3 \cap B|C_1^c) = \frac{1}{2}$. Plug these values to Equation (0.0.5), we have:

$$\begin{aligned} P(C_3|B) &= \frac{0 + \frac{1}{2} \times \frac{2}{3}}{b \times \frac{1}{3} + \frac{1}{2} \times \frac{2}{3}} && (0.0.6) \\ &= \frac{\frac{1}{3}}{\frac{b}{3} + \frac{1}{3}} && (0.0.7) \\ &= \frac{1}{1+b} && (0.0.8) \end{aligned}$$

(iii) Now the second door does not necessarily reveals a goat given door is randomly chosen by the presenter. Given event C_1^c, all the possible arrangements for the second and third doors are $\{$(B,car), (car,B), (N,car), (car,N)$\}$, thus the probability of $P(C_3 \cap G|C_1^c) = P(G|C_1^c) = \frac{1}{2}$. Therefore the probability that the third door conceals the car given you see a goat after the second door is

$$P(C_3|G) = \frac{0 + \frac{1}{2} \times \frac{2}{3}}{1 \times \frac{1}{3} + \frac{1}{2} \times \frac{2}{3}} = \frac{1}{2} \qquad (0.0.9)$$

(iv) Given $p = \frac{2}{3}$ and $\frac{1}{2}$ in (i) and (iii) respectively, which is exactly the bound of the α in (iv). It is reasonable to assume the protocol in (iv) is either (i) with a probability b or (iii) with probability $1-b$. Therefore $p = \frac{2}{3}b + (1-b) \times \frac{1}{2}$. Consider it as the presenter tosses a coin with probability of b for a head. If it is a head, he goes to strategy (i), otherwise (iii). Now it comes to find an expression of b as a function of α such that $p = \alpha$. Solve $\frac{2}{3}b + (1-b) \times \frac{1}{2} = \alpha$ returns $b = 6\alpha - 3$. Therefore we can safely say you never loss by just swapping, but whether you gain depends on the presenter's protocols, as you can start at $p = \frac{1}{2}$ and finish at $p = \frac{1}{2}$ still if $b = 0$.

Question 2

Let F be a distribution function and r a positive integer. Show that the following are distribution functions:

(i) $F(x)^r$,

(ii) $1 - \{1 - F(x)\}^r$

(iii) $F(x) + \{1 - F(x)\} \log\{1 - F(x)\}$

(iv) $(F(x) - 1)e + \exp(1 - F(x))$.

Solution: We can see all these four functions can be written as $g(F(x))$. It is a distribution function only if it is continuous and non-decreasing on $[0, 1]$, with $g(0) = 0$ and $g(1) = 1$ since $0 \le F(x) \le 1$ is a distribution function.

(i) For $g(F(x)) = F(x)^r$, it is continuous. $g(F(x) = 0) = 0$ and $g(F(x) = 1) = 1^r = 1$. Calculate the derivative:

$$\frac{dg}{dx} = \frac{dg}{dF}\frac{dF}{dx} = rF(x)^{r-1}f(x) \geq 0, \tag{0.0.10}$$

where $f(x) \geq 0$ is a pdf of X, given $r > 0$. So $g(F(x))$ is a distribution function. Similar idea applies for (ii)-(iv).

Question 3

Let X be a random variable with distribution function F, and let $a = (a_m : -\infty < m < \infty)$ be a strictly increasing sequence of real numbers satisfying $a_{-m} \to -\infty$ and $a_m \to \infty$ as $m \to \infty$. Define $G(x) = P(X \leq a_m)$ when $a_{-m} \leq x < a_m$, so that G is the distribution function of a discrete random variable. How does the function G behave as the sequence a is chosen in such a way that $\sup_m |a_m - a_{m-1}|$ becomes smaller and smaller?

Solution: By comparing the definitions of $F(x)$ and $G(x)$, it is reasonable to guess that $G(x)$ would approach $F(x)$. Let $\delta = sup_m|a_m - a_{m-1}|$. Note we will only consider the domain $[a_{m-1}, a_m)$ where $G(x)$ is defined. Given

$$\begin{aligned}
F(x) &= P(X \leq x), && (0.0.11)\\
\Rightarrow F(a_{m-1}) &= P(X \leq a_{m-1}) \text{ and} && (0.0.12)\\
F(a_m) &= P(X \leq a_m) && (0.0.13)\\
\Rightarrow F(x) &< F(a_m) && (0.0.14)\\
G(x) &= P(X \leq a_m), && (0.0.15)\\
\Rightarrow F(a_{m-1}) &\leq G(x), \text{since } a \text{ is a strictly increasing sequance} && (0.0.16)
\end{aligned}$$

Note that Line (0.0.14) can also be written as $F(x) \leq F(a_m)$ if $F(x)$ is continuous. Recall that for a continuous random variable, the probability for a single point is zero. We therefore restrict the CDF $F(x)$ as a continuous function. By the definition of δ, we have $\delta \geq |a_m - a_{m-1}|$ for all m. Therefore, $\delta \geq a_m - a_{m-1}$ and $-\delta \leq a_{m-1} - a_m$. Given $a_{m-1} \leq x < a_m$, we have

$$\begin{aligned}
x + \delta \geq a_{m-1} + \delta \geq a_{m-1} + a_m - a_{m-1} = a_m && (0.0.17)\\
x - \delta < a_m - \delta \leq a_m + a_{m-1} - a_m = a_{m-1} && (0.0.18)
\end{aligned}$$

Hence

$$\begin{aligned}
F(x + \delta) &\geq F(a_m) && (0.0.19)\\
F(x - \delta) &\leq F(a_{m-1}) && (0.0.20)
\end{aligned}$$

Note the equivalence for Line (0.0.20) holds when $F(x)$ is continuous. Now the difference between $G(x)$ and $F(x)$, i.e. $|F(x) - G(x)|$.

$$|F(x) - G(x)| \leq |F(a_m) - F(a_{m-1})| \leq |F(x + \delta) - F(x - \delta)|. \tag{0.0.21}$$

According to the question when $\delta \to 0$, $|F(x) - G(x)| \leq 0$ and given $|\cdot| \geq 0$, we have $|F(x) - G(x)| = 0$ when $\delta \to 0$. We therefore conclude that $G(x)$ approaches $F(x)$ for any x at which $F(x)$ is continuous.

Question 4

Which of the following are density functions? Find c and the corresponding distribution function F for those that are.

(a)

$$f(x) = \begin{cases} cx^{-d}, & \text{if } x > 1 \\ 0, & \text{otherwise} \end{cases}$$

(b)

$$f(x) = ce^{x}(1 + e^{x})^{-2}$$

(a) $\int_1^\infty cx^{-d}dx = \frac{c}{1-d}x^{1-d}|_1^\infty$. If $1 - d > 0$, integral goes to infinity. So $d \geq 1$. And $\frac{c}{d-1} = 1$, $c = d - 1$.

(b) Integrate this function and set it to one. We can find that $c = 1$.