Problem Sheet 02

Alexander Pieper (815402)
alexander.pieper@uni-potsdam.de

7. November 2021

Exercise 3

Let X_1, X_2, X_3, X_4 be a sample from $U(0,1)$, and let $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}$ be the order statistic. Determine the density of $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}$.

Answer:

The Probability density function of $Y \sim U(a,b)$ is given by

$$f(t) = \begin{cases} \frac{1}{b-a} & \text{, for } t \in [a,b] \\ 0 & \text{, for } t \notin [a,b] \end{cases} \tag{11}$$

and the cumulative Distribution funciton is given by

$$F(t) = \begin{cases} 0 & \text{, for } t < a \\ \dfrac{t-a}{b-a} & \text{, for } t \in [a,b] \\ 1 & \text{, for } t > a \end{cases} \tag{12}$$

On page 18, equation (10) in the slides of the Lecture 4, we have the density of the order statistic given by

$$f_{X_{(i)}}(t) = \frac{n!}{(i-1)!(n-i)!} f(t) F(t)^{i-1} (1 - F(t))^{n-i} \tag{13}$$

The case for $f(t) = 0$, for $t \notin [a,b]$, will stay the same for all following densities by definition of the Uniform distribution. Therefore we only consider the case $t \in [a,b]$, for $a = 0, b = 1$ which yields $f(t) = 1$ and $F(t) = x$, for $t \in [0,1]$. Also we have $n = 4$, given in the Question.

$$\begin{aligned} f_{X_{(1)}}(t) &= \frac{4!}{0!3!} f(t) F(t)^0 (1 - F(t))^3 && (14) \\ &= \frac{4!}{3!} 1 (1-t)^3 && (15) \\ &= 4(1-t)^3 && (16) \end{aligned}$$

$$\begin{aligned} f_{X_{(2)}}(t) &= \frac{4!}{1!2!} f(t) F(t)^1 (1 - F(t))^2 && (17) \\ &= \frac{4!}{2!} t (1-t)^2 && (18) \\ &= 12t(1-t)^3 && (19) \end{aligned}$$

$$f_{X_{(3)}}(t) = \frac{4!}{2!1!} f(t) F(t)^2 (1 - F(t))^1 \tag{20}$$
$$= \frac{4!}{2!} t^2 (1 - t) \tag{21}$$
$$= 12t^2 (1 - t)^3 \tag{22}$$

$$f_{X_{(4)}}(t) = \frac{4!}{3!0!} f(t) F(t)^3 (1 - F(t))^0 \tag{23}$$
$$= \frac{4!}{3!} t^3 \tag{24}$$
$$= 4t^3 (1 - t)^3 \tag{25}$$

To illustrate these densities, the following plot shows that $X_{(1)}$ has much more mass in the numbers close to 0, whereas $X_{(4)}$ has more mass in the numbers close to 1. This makes sense, because in the order statistc $X_{(4)}$, it is very unlikely to have a low value, because there need to be 3 samples from the same distribution, that are lower than that value and vice versa.

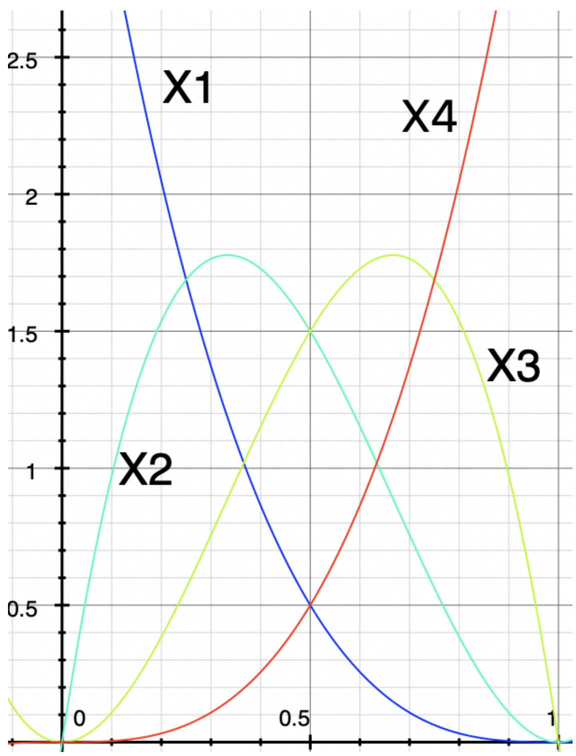

Figure 1: Probability density function of $X_{(1)}$ (blue), $X_{(2)}$ (cyan), $X_{(3)}$ (olive-green), $X_{(4)}$ (red), Graphic created by myself with the App: Grapher