6.3 Order Statistics (Discrete Case)

In the previous sections we transformed or random variable(s) into other random variable(s) through traditional mathematical functions. In this section, we will perform a transformation from many random variables into one random variable. Fortunately, this will not require double, triple or higher order integrals.

Suppose that we roll three dice and get values x_1, x_2 and x_3. We define $x_{(1)}$ as the smallest of our three values. Similarly, we define $x_{(2)}$ as the second smallest and $x_{(3)}$ as the largest of the three.

Example: Three dice are rolled. The values are $x_1 = 5, x_2 = 2$ and $x_3 = 4$. The three order statistics are $x_{(1)} = 2, x_{(2)} = 4$ and $x_{(3)} = 5$.

Example: Four dice are rolled. The values are $x_1 = 5, x_2 = 6, x_3 = 3$ and $x_4 = 3$. Determine the four order statistics.

In the above example, we did not perform any probability calculations. We merely made an attempt to understand what the meaning of this new transformation is. We could have written the transformation in functional notation if we desired. $x_{(1)} = \min\{x_1, x_2, x_3\}$. Consider our first example in section 6.2. We had little interest the value of $U = Y - X$ for specific values of x and y. We had interest in the distribution of U. The same is true here. We want to determine the distribution of our order statistics X_1, X_2 and X_3.

To determine our pmf $f_{X_{(1)}}(x_{(1)})$, we will use the "Brute Force Technique". That is, enumerating everything that can happen and use the Law of Total Probability. Things are actually much easier in the continuous case as we will see in the next section.

Example: In this example, we will be rolling two dice and determining the pmf of the two order statistics $X_{(1)}$ and $X_{(2)}$.

We need to determine $P(X_{(1)} = x_{(1)})$ for $x_{(1)} = 1, 2, 3, 4, 5$ and 6.

$x_{(1)}$	$f(x_{(1)})$
1	11/36
2	9/36
3	7/36
4	5/36
5	3/36
6	1/36

For the minimal order statistic, the easiest value to find would be $f_{X_{(1)}}(6)$. This is because there is only one way for that to occur. Both rolls have to be a 6. So, $f_{X_{(1)}}(6) = 1/36$. For the minimal value to be a 5, we need (5,5), (5,6) or (6,5) giving $f_{X_{(1)}}(5)=3/36$. For the minimal value to be a 4, we need (4,4), (4,5), (4,6), (5,4) or (6,4) giving $f_{X_{(1)}}(4)=5/36$. The rest of the chart is filled out in a similar manner.

$x_{(2)}$	$f(x_{(2)})$
1	
2	
3	
4	
5	
6	

A single die is rolled twice. Determine the maximal order statistic, starting with $f_{X_{(2)}}(1)$ would be easiest.

Example: A random sample of size $n=2$ is taken from a random variable with pmf given below. Determine the pmf of the maximal order statistic $X_{(2)}$. Compare $E[X]$ to $E[X_{(2)}]$

x	1	2	3	4
$f(x)$	.1	.2	.3	.4

Repeat for $X_{(1)}$

Obviously, as the cardinality of the support increases (or the number of items chosen increases) these problems become lengthier. They do not however get more difficult to comprehend.

6.4 Order Statistics (Continuous Case)

In the previous sections we looked at order statistics in the discrete case. Determining the distribution of the order statistics is much easier in the continuous case.

Suppose that $X_1, X_2, \cdots, X_n$ are independent random variable with pdf $f_X(x)$ and CDF $F_X(x)$. Let $X_{(k)}$ denote the kth order statistic. We now determine the pdf of $X_{(k)}$.

Theorem: Let $X_1, X_2, \cdots, X_n$ are independent random variable with pdf $f_X(x)$ and CDF $F_X(x)$. Let $X_{(k)}$ denote the kth order statistic. The pdf of our kth order statistic is:

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} F_X^{k-1}(x) f_X(x) (1 - F_X)^{n-k}$$

Example: Let $X_1, X_2, \cdots, X_8$ be independent random variable with pdf $f_X(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$.

Determine $f_{X_{(3)}}(x)$.