M378K Introduction to Mathematical Statistics

Problem Set #2

Discrete random variables.

2.1. Probability mass function.

Recall the following definition from the last class:

Definition 2.1. *Given a set B, we say that a random variable Y is $B-$*valued *if*

$$\mathbb{P}[Y \in B] = 1.$$

We reserve special terminology for random variables Y depending on the cardinality of the set B from the above definition. In particular, we have the following definition:

Definition 2.2. *A random variable Y is said to be* discrete *if there exists a set S such that :*

- *Y is $S-$valued, and*
- *S is either* **finite** *or* **countable**.

Problem 2.1. *Provide an example of a* **discrete** *random variable.*

Solution: A roll of a fair die.

Our next task is to try to keep track of the probabilities that Y takes specific values from S. In order to be more "economical", we introduce the following concept:

Definition 2.3. *The* support S_Y *of a random variable Y is the* **smallest** *set S such that Y is $S-$valued.*

Problem 2.2. *What is the* **support** *of the random variable you provided as an example in the above problem?*

Solution:

$$S_Y = \{1, 2, 3, 4, 5, 6\}$$

Problem 2.3. *Let $y \in S_Y$ where Y is a discrete random variable. Is it possible to have $\mathbb{P}[Y = y] = 0$?*

Solution: No, it is not possible. If Y is a discrete random variable and there exists $\tilde{y} \in S_Y$ such that $\mathbb{P}[Y = \tilde{y}] = 0$, then we can consider the set $\tilde{S}_Y = S_Y \setminus \{\tilde{y}\}$. Then,

$$\mathbb{P}[Y \in \tilde{S}_Y] = \mathbb{P}[Y \in S_Y] - \mathbb{P}[Y = \tilde{y}] = \mathbb{P}[Y \in S_Y] = 1 \tag{2.1}$$

and $\tilde{S}_Y$ is a smaller set than S_Y such that Y is $\tilde{S}_Y-$valued. This contradicts the definition of support.

Usually, we are interested in calculating and modeling probabilities that look like this

$$\mathbb{P}[Y \in A] \quad \text{for some } A \subseteq S_Y .$$

Note that, if we know the probabilities of the form

$$\mathbb{P}[Y = y] \quad \text{for all } y \in S_Y ,$$

then we can calculate any probability of the above form. *How?*

Solution:

$$\mathbb{P}[Y \in A] = \sum_{y \in A} \mathbb{P}[Y = y] .$$

So, if we "tabulate" the probabilities of the form $\mathbb{P}[Y = y]$ for all $y \in S_Y$, we have sufficient information to calculate any probability of interest to do with the random variable Y. This observation motivates the following definition:

Definition 2.4. *The* probability mass function (pmf) *of a* **discrete** *random variable* Y *is the function* $p_Y : S_Y \to \mathbb{R}$ *defined as*

$$p_Y(y) = \mathbb{P}[Y = y] \quad \textit{for all } y \in S_Y .$$

Can you think of different ways in which to display the pmf?

What is the pmf of the random variable which you provided as an example above?

Solution:

$$p_Y(y) = \begin{cases} \frac{1}{6} & \text{if } y \in \{1,2,3,4,5,6\}, \\ 0 & \text{otherwise.} \end{cases}$$

What are the immediate properties of every pmf?

Solution: For every discrete random variable Y, its pmf p_Y satisfies the following two properties:

- $p_Y(y) > 0$ for all $y \in S_Y$, and
- $\sum_{y \in S_Y} p_Y(y) = 1$.

Does the "reverse" hold, i.e., if a function p_Y satisfies you stated, is it always a pmf of **some** random variable?

Solution: Yes, it does. Given a function p_Y satisfying the two properties above, we can always construct an outcome space Ω, a probability $\mathbb{P}$ on it, and a discrete random variable Y on it such that p_Y is the pmf of Y.

For example, we can consider the outcome space $\Omega = S_Y$. Then, we can define the probability measure $\mathbb{P}$ on Ω as

$$\mathbb{P}[\{\omega\}] = p_Y(\omega) \quad \text{for all } \omega \in \Omega.$$

Finally, we can define the random variable Y as the identity function on Ω, i.e., $Y(\omega) = \omega$ for all $\omega \in \Omega$.

Problem 2.4. *The number of pieces of gossip that break out in a particular high school in a week is modeled by a random variable Y with the following probability mass function:*

$$p_Y(n) = \frac{1}{(n+1)(n+2)} \quad \textit{for all } n \in \mathbb{N}_0.$$

Is the above a well-defined probability mass function?

Solution: We need to verify that the two requirements are satisfied, namely, that

- $p_Y(n) > 0$ for all $n \in \mathbb{N}_0$, and
- $\sum_{n=0}^{\infty} p_Y(n) = 1$.

The first condition is obviously satisfied. As for the second one, we have that for every $N \in \mathbb{N}_0$

$$\sum_{n=0}^{N} p_Y(n) = \sum_{n=0}^{N} \frac{1}{(n+1)(n+2)} = \sum_{n=0}^{N} \left(\frac{1}{n+1} - \frac{1}{n+2} \right) = 1 - \frac{1}{N+2}.$$

As $N \to \infty$, the above sum goes to 1 (which is a definition of the sum of a series).

2.2. **Conditional probability.** In order to "build" more complicated (and useful!) random variables, it helps to review a bit more probability.

Definition 2.5. *Let E and F be two events on the same Ω such that $\mathbb{P}[E] > 0$. The* conditional probability of F given E *is defined as*

$$\mathbb{P}[F \mid E] = \frac{\mathbb{P}[E \cap F]}{\mathbb{P}[E]}\,.$$

Let's spend a moment with the geometric/informational perspective on this definition.

By far, the most popular problems relying on the notion of **conditional probability** are those to do with **specificity** and **sensitivity**[1] of medical tests.

Problem 2.5. *At any given time, 2% of the population actually has a particular disease.*

A test indicates the presence of a particular disease 96% of the time in people who actually have the disease. The same test is positive 1% of the time when actually healthy people are tested.

Calculate the probability that a particular person actually has the disease **given** *that they tested positive.*

Solution: Let F denote the event that a person has the disease and let E denote the event that the person tested positive. We need to calculate the following conditional probability:

$$\mathbb{P}[F \mid E] = \frac{\mathbb{P}[E \cap F]}{\mathbb{P}[E]} = \frac{0.02(0.96)}{0.02(0.96) + 0.98(0.01)} = \frac{2(96)}{2(96) + 98} = \frac{96}{145}\,.$$

[1] https://en.wikipedia.org/wiki/Sensitivity_and_specificity

Moreover, now that we remember the definition of **conditional probability**, we can solve interesting problems such as this one:

Problem 2.6. *The number of pieces of gossip that break out in a particular high school in a week is modeled by a random variable Y with the following probability mass function:*

$$p_Y(n) = \frac{1}{(n+1)(n+2)} \quad \textit{for all } n \in \mathbb{N}_0.$$

Calculate the probability that at least one piece of gossip occurred in a week **given** *that at most four pieces of gossip occurred.*

Solution: Here, we calculate

$$\mathbb{P}[Y \geq 1 \,|\, Y \leq 4] = \frac{\mathbb{P}[1 \leq Y \leq 4]}{\mathbb{P}[Y \leq 4]} = \frac{p_Y(1) + p_Y(2) + p_Y(3) + p_Y(4)}{p_Y(0) + p_Y(1) + p_Y(2) + p_Y(3) + p_Y(4)}\,.$$

Using the same reasoning as in part (ii), we get

$$\mathbb{P}[Y \geq 1 \,|\, Y \leq 4] = \frac{\frac{1}{2} - \frac{1}{6}}{1 - \frac{1}{6}} = \frac{\frac{2}{6}}{\frac{5}{6}} = \frac{2}{5}\,.$$

3. Independent events

What if knowing that an event happened in fact does **not** give any information about the probability of another event?

Definition 3.1. *We say that events E and F on Ω are* independent *if*

$$\mathbb{P}[E \cap F] = \mathbb{P}[E]\mathbb{P}[F].$$

In the case when E or F have a positive probability, it's possible to rewrite the above condition in a different (illustrative!) way. *How?*

Now that we know the notion of **independence**, we can construct random variables in many creative ways.

Example 3.2. *A fair coin is tossed repeatedly and* **independently** *until the first* Heads*. Let the random variable Y represent the total number of* Tails *observed by the end of the procedure.*

What is the support of the random variable Y?

What is the **probability mass function** *of the random variable Y?*