Class Notes 2-6-2019

- **Geometric Distribution Dice Example:**
  What is the chance that it takes more than 10 rolls to roll a six? Let N be the number of rolls until the first six appears.

$$\begin{aligned} P(N>10) &= \sum_{k=11}^{\infty} P(N=k) = \sum_{k=11}^{\infty} p(1-p)^{k-1} \quad \text{where } p = P(\text{ six in one roll}) = 1/6 \\ &= p(1-p)^{10}\sum_{k=0}^{\infty}(1-p)^k = p(1-p)^{10}\frac{1}{1-(1-p)} = (1-p)^{10} \\ &= P(\text{no sixes in the first 10 rolls}) \end{aligned}$$

- **Dice Example:**
  Roll two fair dice until you get a sum of 6 or a sum of 7. What is the chance you get a 7 first? Again this should be viewed in the context of a probability space with an infinite number of trials. We finesse the resulting issues by expressing the event in question as an infinite union of disjoint events, each concerning the outcomes of finite number of rolls, namely

$$A_n = \{\text{no 6 or 7 in the first } n-1 \text{ rolls, and a 7 on the } n^{\text{th}} \text{ roll}\}\ , \quad n = 1, 2, \ldots$$

  The disjoint union $A = \cup_{i=1}^{\infty} A_n$ describes all possible outcomes of interest. For any pair of dice rolls we have

$$P(\text{sum}=7) = \frac{6}{36} \quad \text{and} \quad P(\text{sum}=6) = \frac{5}{36} \quad \text{thus} \quad P(A_n) = \left(1-\frac{11}{36}\right)^{n-1}\frac{6}{36}$$

  and

$$P(A) = \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty}\left(1-\frac{11}{36}\right)^{n-1}\frac{6}{36} = \frac{1}{1-(1-11/36)}\frac{6}{36} = \frac{6}{11} = \frac{6}{5+6}$$

  which intuitively makes sense. Note also that

$$\begin{aligned} P(\text{sum}=7|\text{sum}=6 \cup \text{sum}=7) &= \frac{P(\text{sum}=7)}{P(\text{sum}=6 \cup \text{sum}=7)} = \frac{P(\text{sum}=7)}{P(\text{sum}=6)+P(\text{sum}=7)} \\ &= \frac{6/36}{5/36+6/36} = \frac{6}{5+6} \end{aligned}$$

- **Probability Distributions of Random Variables:**
  An important aspect of any random variable X is its probability distribution which enables us to calculate $P(X \in B)$ for any (reasonable) subset B of the real line $\mathbb{R}$.

- **Probability Mass Function (p.m.f.) for Discrete Random Variables**:
  The distribution of a discrete random variable X with distinct possible values $k_1, k_2, k_3, \ldots$ can be described by its probability mass function (p.m.f.) $p(k) = P(X = k)$ for $k = k_1, k_2, k_3, \ldots$ and $p(k) = 0$ otherwise. If we wish to emphasize the random variable X for which it applies we also write $p_X(k)$.

We must have

$$\sum_k p_X(k) = \sum_k P(X = k) = \sum_{i=1}^{\infty} P(X = k_i) = 1$$

If X has only a finite number N of possible values we replace this by

$$\sum_k p_X(k) = \sum_k P(X = k) = \sum_{i=1}^{N} P(X = k_i) = 1$$

For any $B \subset \mathbb{R}$ we get

$$P(X \in B) = \sum_{k \in B} p_X(k)$$

- **Maximum of Two Fair Dice Rolls:**
  Let X_1, X_2 represent the two fair dice rolls and focus on $X = \max(X_1, X_2)$. X has values $1, 2, \ldots, 6$ with probabilities $1/36, 3/36, 5/36, 7/36, 9/36, 11/36$.

- **Graphical Representation of a pmf:**
  A probability mass function can be represented graphically in various different forms. One possibility is to mark the possible values on the x axis and place vertical rods at those locations (in the y direction) or center vertical boxes at those locations with heights representing the probabilities of the respective values. Others represent the probabilities by the areas of the displayed boxes.

- **Probability Density Functions (p.d.f.) of Random Variables:**
  If $f(x) \geq 0 \;\; \forall x \in \mathbb{R}$ and $\int_{-\infty}^{\infty} f(x)dx = 1$ and if a random variable X satisfies

  $$P(X \leq b) = \int_{-\infty}^{b} f(x)dx \quad \forall b \in \mathbb{R}$$

  then X is called a *continuous* random variable with p.d.f. $f(x) = f_X(x)$.
  Such random variables are not discrete since

  $$P(X = b) = \int_{b}^{b} f(x)dx = 0 \quad \forall b \in \mathbb{R}$$

  For such continuous random variables we thus have

  $$P(X \in [a, b]) = P(X \in (a, b]) = P(X \in [a, b)) = P(X \in (a, b))$$

  For all reasonable subsets $B \subset \mathbb{R}$ we have $\int_B f(x)dx = P(X \in B)$. Such sets B can be any collection of disjoint intervals as long as we can figure out the area under $f(x)$ over all those intervals. See graphical representation in Fig. 3.1 (text)
  If we change $f(x)$ at a finite or countably infinite number of points that will not change $P(X \in B)$, i.e., we still have the same probability distribution. The term continuous random variable is a bit of a misnomer, return to this later when introducing distribution functions.

- **Potential Density Functions(??)**
  We use indicator functions such as $I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ if $x \notin A$.

  $$f_1(x) = \frac{1}{x^2} I_{\{x \geq 1\}}(x), \quad f_2(x) = b\sqrt{a^2 - x^2} I_{[0,a]}(|x|), \text{for } a, b > 0, \quad f_3(x) = c \sin x \; I_{[0,2\pi]}(x)$$

f_1 is a density function since $f_1 \geq 0 \;\; \forall x \in \mathbb{R}$ and

$$\int_{-\infty}^{\infty} f_1(x)\, dx = \int_1^{\infty} \frac{1}{x^2}\, dx = \left(-\frac{1}{x}\right)\Big|_1^{\infty} = 1$$

Since $y = \sqrt{a^2 - x^2}$ represents the y ordinate of a point on a circle with radius a at the abscissa x $(y^2 + x^2 = a^2)$ we have

$$\int_{-\infty}^{\infty} f_2(x)\, dx = b \int_{-a}^{a} \sqrt{a^2 - x^2}\, dx = b\frac{1}{2}\pi a^2$$

which is 1 exactly when $b = 2/(\pi a^2)$, i.e., f_2 could be a p.d.f. However, f_3 cannot be a density for the simple reason that it switches sign for any $c \neq 0$ and is identically 0 when $c = 0$.

- **Uniform Distributions or Random Variables:**
  Let $[a, b]$ $(a < b)$ be any finite interval on the real line $\mathbb{R}$. A random variable X has the uniform distribution on the interval $[a, b]$ if X has density

$$f(x) = \frac{1}{b-a}\, I_{[a,b]}(x)$$

  for short we write $X \sim \text{Unif}[a, b]$ or also $X \sim \text{U}[a, b]$, or also $X \sim \text{U}(a, b)$.
  If $[c, d] \subset [a, b]$ then

$$P(X \in [c, d]) = P(c \leq X \leq d) = \int_c^d \frac{1}{b-a} dx = \frac{d-c}{b-a}$$

- **Uniform Example:**
  Let $Y \sim \text{U}(-2, 5)$, what is $P(|Y| \geq 1.5)$?

$$\begin{aligned} P(|Y| \geq 1.5) &= P(Y \in [-2, -1.5] \cup [1.5, 5]) = P(-2 \leq Y \leq -1.5) + P(1.5 \leq X \leq 5) \\ &= \frac{(-1.5) - (-2)}{5 - (-2)} + \frac{5 - 1.5}{5 - (-2)} = \frac{.5}{7} + \frac{3.5}{7} = \frac{4}{7} \end{aligned}$$

- **Infinitesimals or Probability Interpretation of Densities:**
  $f(x)$ is not the probability $P(X = x) = 0$. However, when f is continuous at a point a then

$$P(a - \epsilon/2 < X < a + \epsilon/2) \approx \epsilon f(a) \quad \text{for small } \epsilon > 0$$

  because

$$P(a - \epsilon/2 < X < a + \epsilon/2) = \int_{a-\epsilon/2}^{a-\epsilon/2} f(x)\, dx \approx \epsilon f(a)$$

  If f is right continuous at a we have

$$P(a < X < a + \epsilon) \approx \epsilon f(a)$$

  with a corresponding statement when left continuity holds at a.