Today: § 6.1-6.2
Next: § 6.3, 8.1-8.3

HW 6 due **TONIGHT** by 11:59pm
Lab 6 due **Friday, Nov 21**, by 11:59pm

Midterm 2: Next Wednesday covering Chapters 3-5.
Practice midterms posted this weekend.
Definition: Given (discrete) random variables $X_1, X_2, \ldots, X_n$ all defined on the same sample space, their joint distribution is the collection of all

$$\begin{align*}
\{ & \Pr(X_1 = k_1, X_2 = k_2, \ldots, X_n = k_n) \\
& \text{all possible values } k_i \text{ of } X_1, k_i \text{ of } X_2, \ldots, k_n \text{ of } X_n \}
\end{align*}$$

E.g., $X, Y \sim \text{Ber}(p)$,

1. $X = Y$

2. $X, Y$ independent

$$\begin{array}{c|cc}
& 0 & 1 \\
\hline
0 & 1-p & (1-p)^2 \\
1-p & 0 & p(1-p) \\
\hline
(1-p)p & \text{p} & p^2
\end{array}$$

$$p_{X,Y}(k_1, k_2) = \Pr(X = k_1, Y = k_2)$$
Recovering $X_j$ from $X = (X_1, X_2, \ldots, X_n)$: Marginals

Suppose we know $p_X(k)$ for all $k = (k_1, k_2, \ldots, k_n)$.

How can we find $p_{X_1}(t)$?

Eg. Toss a fair coin twice. $X_1, X_2 \in \{0, 1\}$

$$P(X_1 = 0) = P(X_1 = 0, X_2 = 0) + P(X_1 = 0, X_2 = 1)$$

$$\{X_1 = 0\} = \{X_1 = 0, X_2 = 0\} \cup \{X_1 = 0, X_2 = 1\}$$

In general,

$$P(X_1 = t) = \sum_{k_2, k_3, \ldots, k_n} P(X_1 = t, X_2 = k_2, \ldots, X_n = k_n) = \sum_{k_2, k_3, \ldots, k_n} p_X(t, k_2, k_3, \ldots, k_n)$$

$$\{X_1 = t\} = \bigcup_{k_2, k_3, \ldots, k_n} \{X_1 = t, X_2 = k_2, X_3 = k_3, \ldots, X_n = k_n\}$$

$$P(X = t) = \sum_k p_{(X,Y)}(t, k)$$

$$P(Y = t) = \sum_k p_{X|Y}(k, t)$$
Eg. Toss a fair coin 3 times. \( X = \# \text{tails in first toss (0 or 1)} \)
\( Y = \text{total \#tails in all 3 (0, 1, 2, 3)} \)

- \( X = 0 \) \( \frac{1}{8} \) \( \frac{1}{4} \) \( \frac{1}{8} \) \( 0 \) \( \frac{1}{2} \)
- \( X = 1 \) \( \frac{1}{8} \) \( \frac{1}{4} \) \( \frac{1}{8} \) \( 0 \) \( \frac{1}{2} \)

\( Y \) \( \text{Ber}(\frac{1}{2}) \)

Outcome:
- HHH
- HHT
- HTT
- TTH
- TTT

\( X \) \( Y \)
- 0    0
- 0    1
- 0    2
- 1    1
- 1    2
- 1    3

Question: Are \( X, Y \) independent?

\( P(X=0, Y=0) = 0 \)

\( P(X=1)P(Y=0) = \frac{1}{2} \cdot \frac{1}{8} = \frac{1}{16} \neq 0 \)

No!
Joint distributions are just distributions of \textbf{random vectors}.

\[ X_1, X_2, \ldots, X_n \xrightarrow{\text{d}} X = (X_1, X_2, \ldots, X_n) \]

possible values for \( X \) are vectors \((k_1, \ldots, k_n)\).

\textbf{Eg. Multinomial Distribution.}

Often, trials have more than 2 outcomes.

Consider a trial with \( r \) possible outcomes, \( \omega \) probabilities \( p_1, p_2, \ldots, p_r \)

\[ p_1 + p_2 + \cdots + p_r = 1 \]

Perform \( n \) trials.

For \( 1 \leq j \leq r \), \( X_j \) = \#times we get outcome \( j \).

Possible values for \( X = (X_1, \ldots, X_r) \) are \((k_1, \ldots, k_r) \subseteq \{0, 1, \ldots, n\}^r \)

\[ k_1 + \cdots + k_r = n \]

\[
\mathbb{P}(X = k) = \frac{n!}{k_1! k_2! \cdots k_r!} \]

\text{pmf of Mult}(n; p_1, \ldots, p_r) \Rightarrow \binom{n}{k_1, k_2, \ldots, k_r} = \frac{n!}{k_1! k_2! \cdots k_r!}

\]
Sample 10 times with replacement:

\[ P(3 \text{ green}, 2 \text{ red}, 5 \text{ blue}) \]

\[ G = \# \text{green}, \quad R = \# \text{red}, \quad B = \# \text{blues} \]

\[ X = (G, R, B) \sim \text{Mult}(10; \frac{1}{6}, \frac{1}{3}, \frac{1}{2}) \]

\[ P(X = (3, 2, 5)) = \frac{10!}{3!2!5!} \left( \frac{1}{6} \right)^3 \left( \frac{1}{3} \right)^2 \left( \frac{1}{2} \right)^5 \approx 4.05\% \]

**Note:** if \( r = 2 \), \( \text{Mult}(n; p, q) \)

**Eg.** Suppose \( X \sim \text{Mult}(n; p_1, p_2, \ldots, p_r) \). Find the marginal distribution of \( X_1 \).

\[ P(X_1 = t) = \sum_{t, k_2, \ldots, k_r \geq 0} \frac{n!}{t! k_2! \ldots k_r!} p_1^t p_2^{k_2} \cdots p_r^{k_r} \]

\[ \Rightarrow \sum_{t+k_2+\cdots+k_r=n} X_1 = \# \text{successes in } n \text{ indep. trials} \]

\[ \text{were success = outcome 1 (} IP = p_1\text{), failure = outcomes } 2-r \text{ (} IP = p_2+\cdots+p_r\text{)} \]

\[ X_1 \sim \text{Bin}(n, p) \]
Expectations

Let \( \mathbf{X} = (X_1, \ldots, X_n) \) be a (discrete) random vector with joint probability mass function \( p_X \).

If \( g: \mathbb{R}^n \to \mathbb{R} \) is a function, \( Y = g(\mathbf{X}) \) is a random variable (still discrete).

\[
E(Y) = \sum_t t \cdot p(Y = t)
\]

\[
= \sum_k g(k) \cdot p(X = k)
\]

\[
= \sum_k g(k) \cdot p_X(k)
\]

Eg. Toss a fair coin twice, \( X_1, X_2 \in \{0, 1\} \).

\[
E(X_1 X_2) = \sum_{k_1, k_2} g(k_1, k_2) \cdot p(X_1 = k_1, X_2 = k_2)
\]

\[
= \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 0
\]

\[
= \frac{1}{2} = \mathbb{E}(X) \mathbb{E}(Y)
\]

\[
\mathbb{E}(X) = \frac{1}{2}
\]

\[
\mathbb{E}(Y) = \frac{1}{2}
\]

\[
\mathbb{E}(X_1 X_2) = \frac{1}{4}
\]
Jointly Continuous Random Vectors

A random vector \( \mathbf{X} = (X_1, \ldots, X_n) \) has a pdf \( f_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{R}_+ \) if, for "nice" subsets \( B \subseteq \mathbb{R}^n \)

\[
P(\mathbf{X} \in B) = \int_B f_{\mathbf{X}}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n.
\]

(we say \( X_1, X_2, \ldots, X_n \) are jointly continuous.)

Properties:
1. \( f_{\mathbf{X}} \geq 0 \)
2. \( \int_{\mathbb{R}^n} f_{\mathbf{X}} \, dx = 1 \).

E.g. Standard Multivariate Normal

\[
f(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2} \sum_{i=1}^n x_i^2}
\]

\[
\int_{\mathbb{R}^n} f(x_1, \ldots, x_n) \, dx_1 \cdots dx_n = (2\pi)^{-n/2} \int_{-\infty}^{\infty} e^{-x_1^2/2} \, dx_1 \cdots \int_{-\infty}^{\infty} e^{-x_n^2/2} \, dx_n = (2\pi)^{-n/2} (\sqrt{2\pi}) \cdots (\sqrt{2\pi}) = 1
\]