Section 8
CSE312 23Wi

Solutions by Anna Karlin, William Howard-Snyder and Aleks Jovicic
Slides: Matthew Shang and Alex(Xinlei) Liu
Administrative

- Pset 6 due yesterday
- Pset 7 due **Wednesday March 1**
### Review: Joint Distributions

#### 1) Multivariate: Discrete to Continuous:

<table>
<thead>
<tr>
<th></th>
<th>Discrete</th>
<th>Continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Joint PMF/PDF</strong></td>
<td>$p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y)$</td>
<td>$f_{X,Y}(x,y) \neq \mathbb{P}(X = x, Y = y)$</td>
</tr>
<tr>
<td><strong>Joint range/support</strong></td>
<td>${(x,y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x,y) &gt; 0}$</td>
<td>${(x,y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x,y) &gt; 0}$</td>
</tr>
<tr>
<td><strong>Joint CDF</strong></td>
<td>$F_{X,Y}(x,y) = \sum_{t \leq x, s \leq y} p_{X,Y}(t,s)$</td>
<td>$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s) , ds , dt$</td>
</tr>
<tr>
<td><strong>Normalization</strong></td>
<td>$\sum_{x,y} p_{X,Y}(x,y) = 1$</td>
<td>$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) , dx , dy = 1$</td>
</tr>
<tr>
<td><strong>Marginal PMF/PDF</strong></td>
<td>$p_X(x) = \sum_y p_{X,Y}(x,y)$</td>
<td>$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) , dy$</td>
</tr>
<tr>
<td><strong>Expectation</strong></td>
<td>$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y) p_{X,Y}(x,y)$</td>
<td>$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) , dx , dy$</td>
</tr>
<tr>
<td><strong>Independence</strong></td>
<td>$\forall x,y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$</td>
<td>$\forall x,y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$</td>
</tr>
<tr>
<td></td>
<td>$\Omega_{X,Y} = \Omega_X \times \Omega_Y$</td>
<td>$\Omega_{X,Y} = \Omega_X \times \Omega_Y$</td>
</tr>
<tr>
<td><strong>Conditional PMF/PDF</strong></td>
<td>$p_{X</td>
<td>Y}(x</td>
</tr>
<tr>
<td><strong>Conditional Expectation</strong></td>
<td>$\mathbb{E}[X</td>
<td>Y = y] = \sum_x x \cdot p_{X</td>
</tr>
</tbody>
</table>
3) **Central Limit Theorem (CLT):** Let $X_1, \ldots, X_n$ be iid random variables with $\mathbb{E}[X_i] = \mu$ and $Var(X_i) = \sigma^2$. Let $X = \sum_{i=1}^{n} X_i$, which has $\mathbb{E}[X] = n\mu$ and $Var(X) = n\sigma^2$. Let $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, which has $\mathbb{E}[\bar{X}] = \mu$ and $Var(\bar{X}) = \frac{\sigma^2}{n}$. $\bar{X}$ is called the *sample mean*. Then, as $n \to \infty$, $\bar{X}$ approaches the normal distribution $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Standardizing, this is equivalent to $Y = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ approaching $\mathcal{N}(0, 1)$. Similarly, as $n \to \infty$, $X$ approaches $\mathcal{N}(n\mu, n\sigma^2)$ and $Y' = \frac{X - n\mu}{\sigma\sqrt{n}}$ approaches $\mathcal{N}(0, 1)$.

It is no surprise that $\bar{X}$ has mean $\mu$ and variance $\sigma^2/n$ – this can be done with simple calculations. The importance of the CLT is that, for large $n$, regardless of what distribution $X_i$ comes from, $\bar{X}$ is *approximately normally distributed with mean $\mu$ and variance $\sigma^2/n$*. Don't forget the continuity correction, only when $X_1, \ldots, X_n$ are discrete random variables.

Here is the **Standard normal table.**
Problem 1: Joint PMFs

Suppose $X$ and $Y$ have the following joint PMF:

<table>
<thead>
<tr>
<th>X/Y</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
<td>0</td>
<td>0.4</td>
</tr>
</tbody>
</table>

a) Identify the range of $X$ ($\Omega_X$), the range of $Y$ ($\Omega_Y$), and their joint range ($\Omega_{X,Y}$).

b) Find the marginal PMF for $X$, $p_X(x)$ for $x \in \Omega_X$.

c) Find the marginal PMF for $Y$, $p_Y(y)$ for $y \in \Omega_Y$.

d) Are $X$ and $Y$ independent? Why or why not?

e) Find $E[X^3Y]$. 
Problem 1 Solution: a)

a) Identify the range of $X$ ($\Omega_X$), the range of $Y$ ($\Omega_Y$), and their joint range ($\Omega_{X,Y}$).

$\Omega_X = \{0, 1\}$, $\Omega_Y = \{1, 2, 3\}$, and $\Omega_{X,Y} = \{(0, 2), (0, 3), (1, 1), (1, 3)\}$
Problem 1 Solution: b)

b) Find the marginal PMF for $X$, $p_X(x)$ for $x \in \Omega_X$.

\[
p_X(0) = \sum_y p_{X,Y}(0,y) = 0 + 0.2 + 0.1 = 0.3
\]

\[
p_X(1) = 1 - p_X(0) = 0.7
\]
Problem 1 Solution: c)

c) Find the marginal PMF for $Y$, $p_Y(y)$ for $y \in \Omega_Y$.

\[
p_Y(1) = \sum_x p_{X,Y}(x, 1) = 0 + 0.3 = 0.3
\]

\[
p_Y(2) = \sum_x p_{X,Y}(x, 2) = 0.2 + 0 = 0.2
\]

\[
p_Y(3) = \sum_x p_{X,Y}(x, 3) = 0.1 + 0.4 = 0.5
\]
Problem 1 Solution: d)

d) Are $X$ and $Y$ independent? Why or why not?

No, since a necessary condition is that $\Omega_{X,Y} = \Omega_X \times \Omega_Y$. 
Problem 1 Solution: e)

e) Find $\mathbb{E}[X^3Y]$.

Note that $X^3 = X$ since $X$ takes values in $\{0, 1\}$.

$$\mathbb{E}[X^3Y] = \mathbb{E}[XY] = \sum_{(x,y) \in \Omega_{X,Y}} xy p_{X,Y}(x,y) = 1 \cdot 1 \cdot 0.3 + 1 \cdot 3 \cdot 0.4 = 1.5$$
Problem 9: Confidence Intervals

Suppose that $X_1, \ldots, X_n$ are i.i.d. samples from a normal distribution with unknown mean $\mu$ and variance 36. How big does $n$ need to be so that $\mu$ is in 

$$[\bar{X} - 0.11, \bar{X} + 0.11]$$

with probability at least 0.97?

Recall that 

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$ 

You may use the fact that $\Phi^{-1}(0.985) = 2.17$. 
Problem 9 Solution

Our goal is to find $n$ such that $\mu$ lies within 0.11 of $\bar{X}$ 97% of the time. This is equivalent to finding $n$ such that the probability that $\mu$ lies outside the range is less than 3%.

$$\mathbb{P}(|\bar{X} - \mu| > 0.11) \leq 0.03$$
Problem 9 Solution

Let us define $Z = \frac{\bar{X} - \mu}{\sigma}$. We can solve for $\sigma$ by using the Properties of Variance. Since

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

we can say that

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right)$$

Using the Properties of Variance and the fact that $X_i$'s are i.i.d., $\text{Var}(\bar{X}) = \frac{1}{n^2} \cdot n \cdot 36 = \frac{36}{n}$, so $\sigma = \frac{6}{\sqrt{n}}$. 
Problem 9 Solution

\[ P(|\bar{X} - \mu| > 0.11) \leq 0.03 \]
\[ P(|Z| \cdot \sigma > 0.11) \leq 0.03 \]

[Definition of Z]
Problem 9 Solution

\[ P(\bar{X} - \mu > 0.11) \leq 0.03 \]
\[ P(|Z| \cdot \sigma > 0.11) \leq 0.03 \]
\[ P \left( |Z| > \frac{0.11}{6} \sqrt{n} \right) \leq 0.03 \]
\[ P \left( Z < -\frac{0.11}{6} \sqrt{n} \right) \leq 0.015 \]
\[ \Phi \left( -\frac{0.11}{6} \sqrt{n} \right) \leq 0.015 \]

[Definition of Z]

[Symmetry of Normal Dist.]

[CDF of Standard Norm.]
Therefore the final answer should be 14010.
Problem 6: Continuous Joint Density

The joint density of $X$ and $Y$ is given by

$$f_{X,Y}(x, y) = \begin{cases} xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

and the joint density of $W$ and $V$ is given by

$$f_{W,V}(w, v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are $X$ and $Y$ independent? Are $W$ and $V$ independent?
Problem 6 Solution

For two random variables $X, Y$ to be independent, we must have $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ for all $x \in \Omega_X, \ y \in \Omega_Y$. Let’s start with $X$ and $Y$ by finding their marginal PDFs. By definition, and using the fact that the joint PDF is 0 outside of $y > 0$, we get:

$$f_X(x) = \int_0^\infty xe^{-(x+y)}\,dy = e^{-x}x$$
Problem 6 Solution

We do the same to get the PDF of \( Y \), again over the range \( x > 0 \):

\[
f_Y(y) = \int_0^\infty xe^{-(x+y)} \, dx = e^{-y}
\]

Since \( e^{-x}x \cdot e^{-y} = xe^{-x-y} = xe^{-(x+y)} \) for all \( x, y > 0 \), \( X \) and \( Y \) are independent.
Problem 6 Solution

We can see that $W$ and $V$ are not independent simply by observing that $\Omega_W = (0, 1)$ and $\Omega_V = (0, 1)$, but $\Omega_{W,V}$ is not equal to their Cartesian product. Specifically, looking at their range of $f_{W,V}(w,v)$. Graphing it with $w$ as the "x-axis" and $v$ as the "y-axis", we see that: