Joint Distributions

H. Krieger, Mathematics 62, Harvey Mudd College

Fall, 2008

Definition 1: Let $X$ and $Y$ be discrete random variables defined on the same sample space. Then the joint probability mass function or joint pmf is defined by

$$p(x, y) = P(X = x, Y = y).$$

Note that $p(x, y) \geq 0$ and $\sum_{(x,y)} p(x, y) = 1$. In this case, if $I = (a, b]$ and $J = (c, d]$ are intervals of real numbers, then

$$P(\{X \in I\} \cap \{Y \in J\}) = \sum_{x \in I} \sum_{y \in J} p(x, y).$$

Furthermore, we see that the marginal distributions of $X$ and $Y$ can be described by the pmf’s given by

$$p_X(x) = P(X = x) = \sum_y p(x, y) \quad \text{and} \quad p_Y(y) = P(Y = y) = \sum_x p(x, y).$$

With this notation, $X$ and $Y$ are independent random variables iff for every pair $(x, y)$ we have

$$p(x, y) = p_X(x)p_Y(y).$$

If $\phi$ is a real valued function defined for values of $(x, y)$ which lie in the range of $(X, Y)$, then we can show that

$$E(\phi(X, Y)) = \sum_{(x,y)} \phi(x, y)p(x, y).$$

In particular, letting $\phi(x, y) = ax + by$, we see that expected value is linear, namely:

$$E(aX + bY) = \sum_{(x,y)} (ax + by)p(x, y) = aE(X) + bE(Y).$$

Definition 2: Now, let $X$ and $Y$ be continuous random variables defined on the same sample space. Then the joint probability density function or joint
pdf is defined to be a function $f(x, y) \geq 0$ with $\int_{\mathbb{R}^2} f(x, y) \, dx \, dy = 1$. In this case, if $I = (a, b]$ and $J = (c, d]$ are intervals of real numbers, then

$$P(\{X \in I\} \cap \{Y \in J\}) = \int_c^d \int_a^b f(x, y) \, dx \, dy.$$ 

Furthermore, we see that the marginal distributions of $X$ and $Y$ can be described by the pdf’s given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx.$$ 

We say that the continuous random variables $X$ and $Y$ are independent iff for every pair $(x, y)$ we have

$$f(x, y) = f_X(x)f_Y(y).$$

If $\phi$ is a real valued function defined for values of $(x, y)$ which lie in the range of $(X, Y)$, then we can show that

$$E(\phi(X, Y)) = \int_{\mathbb{R}^2} \phi(x, y) f(x, y) \, dx \, dy.$$ 

Once again, letting $\phi(x, y) = ax + by$, we see that expected value is linear in this case as well, namely:

$$E(aX + bY) = \int_{\mathbb{R}^2} (ax + by) f(x, y) \, dx \, dy = aE(X) + bE(Y).$$

**Some Important Consequences**

1. If $X$ and $Y$ are random variables such that $X \leq Y$ with probability one, then $E(X) \leq E(Y)$.

2. Chebyshev’s inequality holds, namely:

$$P(|X - \mu_X| \geq k\sigma_X) \leq 1/k^2.$$ 

3. If we define the covariance of random variables $X$ and $Y$ to be

$$\text{Cov}(X, Y) = E([X - \mu_X][Y - \mu_Y]) = E(XY) - \mu_X\mu_Y,$$

we see that

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).$$

Consequently, defining two random variables $X$ and $Y$ to be uncorrelated iff $\text{Cov}(X, Y) = 0$, we see that the following are equivalent statements:

(a) $X$ and $Y$ are uncorrelated.
(b) $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$.

(c) $E(XY) = E(X)E(Y)$.

4. If $X$ and $Y$ are independent, then $E(XY) = E(X)E(Y)$ and, hence, $X$ and $Y$ are uncorrelated. The converse is not true!

5. We can define the correlation coefficient of two random variables $X$ and $Y$ to be

\[ \rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}. \]

The Cauchy-Schwarz inequality shows that $|\text{Cov}(X,Y)| \leq \sigma_X \sigma_Y$ and hence that $-1 \leq \rho(X,Y) \leq +1$.

Examples: Consider the following examples of discrete joint distributions,

A: Dependent but uncorrelated

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y$</th>
<th>$P(Y/X)$</th>
<th>$P(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/4</td>
<td>1</td>
<td>0</td>
<td>1/8</td>
</tr>
<tr>
<td>3/8</td>
<td>0</td>
<td>1/16</td>
<td>1/4</td>
</tr>
<tr>
<td>3/8</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

B: Dependent and highly correlated

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y$</th>
<th>$P(Y/X)$</th>
<th>$P(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/4</td>
<td>1</td>
<td>0</td>
<td>1/4</td>
</tr>
<tr>
<td>1/2</td>
<td>1</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>1/4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

C: Independent and, hence, uncorrelated

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y$</th>
<th>$P(Y/X)$</th>
<th>$P(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/4</td>
<td>1</td>
<td>1/16</td>
<td>1/8</td>
</tr>
<tr>
<td>1/2</td>
<td>1</td>
<td>1/16</td>
<td>1/4</td>
</tr>
<tr>
<td>1/4</td>
<td>0</td>
<td>1/16</td>
<td>1/8</td>
</tr>
</tbody>
</table>

Exercises:

1. Verify the descriptions given for each of these examples. In particular, show that for Example B, the correlation coefficient $\rho(X,Y) = -1$. Hints: What does it take to show that two random variables are dependent? Show $X$ and $Y$ are uncorrelated by showing that $E(XY) = E(X)E(Y)$. 

3
2. Look at the marginal distributions in Examples B and C. Suppose $X_B$ and $X_C$ are random variables with the same distribution. Also, suppose $Y_B$ and $Y_C$ are random variables with the same distribution. What can we say about the joint distribution of $X_B$ and $Y_B$ as compared to the joint distribution of $X_C$ and $Y_C$?

3. Show that for Examples B and C, the marginal distributions of both $X$ and $Y$ are binomial($n = 2, p = 1/2$) and, hence, that $Var(X) = Var(Y) = npq = 2 \cdot 1/2 \cdot 1/2 = 1/2$.

4. For Example A, explain (without a complete calculation) why $Var(X) < 1/2 < Var(Y)$. Verify this by making the calculation.

5. For each of the 3 examples, find the pmf for the random variable $X + Y$. You should recognize these distributions! Does knowing the distribution of $X + Y$ determine the joint distribution of $X$ and $Y$?