Bivariate Distributions

**Definition:** class of multivariate probability distributions describing joint variation of outcomes of two random variables (discrete or continuous), e.g., temperature and precipitation, household size and number of cars owned.

**Example:** survey results of 100 households reporting # of people per household (household size) and # of cars owned per household:

<table>
<thead>
<tr>
<th>size (x)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>tot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>10</td>
<td>6</td>
<td>3</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>12</td>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>tot.</td>
<td>22</td>
<td>25</td>
<td>27</td>
<td>26</td>
<td>100</td>
</tr>
</tbody>
</table>

**Bivariate PMF:** probability mass function \( f_{X,Y}(x, y) \) that two RVs \( X \) and \( Y \) attain simultaneously two respective outcomes \( x \) and \( y \):

\[
 f_{X,Y}(x, y) = \text{Prob}(X = x, Y = y) = \text{number of samples occurring jointly in classes } x \text{ and } y / \text{total number of samples} 
\]

<table>
<thead>
<tr>
<th>size (x)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>( f_X(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.10</td>
<td>0.08</td>
<td>0.03</td>
<td>0.02</td>
<td>0.23</td>
</tr>
<tr>
<td>3</td>
<td>0.07</td>
<td>0.10</td>
<td>0.06</td>
<td>0.03</td>
<td>0.26</td>
</tr>
<tr>
<td>4</td>
<td>0.04</td>
<td>0.05</td>
<td>0.12</td>
<td>0.06</td>
<td>0.27</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>0.02</td>
<td>0.06</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td>( f_Y(y) )</td>
<td>0.22</td>
<td>0.25</td>
<td>0.27</td>
<td>0.26</td>
<td>1.00</td>
</tr>
</tbody>
</table>

note that \( 0 \leq f_{X,Y}(x, y) \leq 1 \), and \( \sum_x \sum_y f_{X,Y}(x, y) = 1 \)
### Discrete Marginal Distributions

<table>
<thead>
<tr>
<th># of cars owned (y)</th>
<th>size (x)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>( f_X(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>size (x)</td>
<td></td>
<td>0.10</td>
<td>0.08</td>
<td>0.03</td>
<td>0.02</td>
<td>0.23</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.07</td>
<td>0.10</td>
<td>0.06</td>
<td>0.03</td>
<td>0.26</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.04</td>
<td>0.05</td>
<td>0.12</td>
<td>0.06</td>
<td>0.27</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.01</td>
<td>0.02</td>
<td>0.06</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td>( f_Y(y) )</td>
<td></td>
<td>0.22</td>
<td>0.25</td>
<td>0.27</td>
<td>0.26</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**Marginal (univariate) distributions:**
- univariate PMFs of the two RVs \( X \) and \( Y \)
- obtained by integrating (summing up) the columns and rows of the PMF table:

\[
 f_X(x) = \sum_y f_{X,Y}(x,y) \quad \text{and} \quad f_Y(y) = \sum_x f_{X,Y}(x,y)
\]

\[
 f_X(x) = \text{Prob}\{X = x, Y \in [\pm \infty] \}
\]

marginal PMF = bivariate PMF with one RV unconstrained

**Examples:**
- probability of a randomly selected household with 3 members:
  \[
  f_X(3) = 0.07 + 0.10 + 0.06 + 0.03 = 0.26
  \]
- probability of a randomly selected household with no cars:
  \[
  f_Y(0) = 0.10 + 0.07 + 0.04 + 0.01 = 0.22
  \]
Covariance and Correlation

<table>
<thead>
<tr>
<th># of cars owned (y)</th>
<th>size (x)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>$f_X(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>0.10</td>
<td>0.08</td>
<td>0.03</td>
<td>0.02</td>
<td>0.23</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.07</td>
<td>0.10</td>
<td>0.06</td>
<td>0.03</td>
<td>0.26</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.04</td>
<td>0.05</td>
<td>0.12</td>
<td>0.06</td>
<td>0.27</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.01</td>
<td>0.02</td>
<td>0.06</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td>$f_Y(y)$</td>
<td></td>
<td>0.22</td>
<td>0.25</td>
<td>0.27</td>
<td>0.26</td>
<td>1.00</td>
</tr>
</tbody>
</table>

$$\sigma_{XY} = Cov\{X, Y\} = E[[X - E\{X\}][Y - E\{Y\}]] = E\{XY\} - E\{X\}E\{Y\}$$

Covariance calculation:

$$\sigma_{XY} = \sum_{x=2}^{5} \sum_{y=0}^{3} xyf_{X,Y}(x, y) - \sum_{x=2}^{5} xf_X(x) \sum_{y=0}^{3} yf_Y(y)$$

$$= \frac{6.16 - 3.52 \times 1.57}{0.6336}$$

Correlation coefficient: standardized (unit-free) version of covariance:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \in [-1, +1]$$

$\sigma_X$ = variance of RV $X$; $\sigma_Y$ = variance of RV $Y$

Correlation coefficient calculation:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sqrt{\sum_{x=2}^{5} x^2f_X(x) - E\{X\}^2} \sqrt{\sum_{y=0}^{3} y^2f_Y(y) - E\{Y\}^2}}$$

$$= \frac{0.6336}{1.0907 \times 1.0978} = 0.5292$$

Correlation versus Dependence

<table>
<thead>
<tr>
<th># of cars owned (y)</th>
<th>size (x)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>$f_X(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>0.10</td>
<td>0.08</td>
<td>0.03</td>
<td>0.02</td>
<td>0.23</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.07</td>
<td>0.10</td>
<td>0.06</td>
<td>0.03</td>
<td>0.26</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.04</td>
<td>0.05</td>
<td>0.12</td>
<td>0.06</td>
<td>0.27</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.01</td>
<td>0.02</td>
<td>0.06</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td>$f_Y(y)$</td>
<td></td>
<td>0.22</td>
<td>0.25</td>
<td>0.27</td>
<td>0.26</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Independence:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) \ \forall x, y$$

bivariate PMF = product of two marginal PMFs, for all possible combinations of $x$ and $y$

Example:

$$f_{X,Y}(3, 1) = 0.10 \neq 0.26 \times 0.25 = f_X(3)f_Y(1)$$

Important note:

- lack of correlation, i.e., $\sigma_{XY} = \rho_{XY} = 0$, does not entail independence, because possibly $f_{X,Y}(x, y) \neq f_X(x)f_Y(y)$ for some combinations of $x$ and $y$

- independence, however, entails lack of correlation:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y), \ \forall x, y \Rightarrow \sigma_{XY} = \rho_{XY} = 0$$

Note: Covariance $\sigma_{XY}$ and correlation $\rho_{XY}$ are just a moments, i.e., summary statistics, of the bivariate PMF $f_{X,Y}(x, y)$
**Joint Variability of Two Continuous Attributes**

Consider two random variables \( Z \) and \( Y \), and their sample scatter plot, i.e., \( N \) realized pairs \( \{(z_n, y_n), n = 1, \ldots, N\} \)

**Sample Bivariate Density Histogram**

**Discretization step:** mesh of rectangular bins overlaid on sample scatter plot

**Normalized sample frequencies:** = # of pairs falling in each bin, divided by: (i) total # of pairs, and (ii) bin area

**Joint and marginal PDFs:**

\[
(Z, Y) \sim f_{ZY}(z, y) = \frac{\text{Prob}\{Z \in z \pm dz, Y \in y \pm dy\}}{\text{bivariate PDF}}
\]

\[
Z \sim f_Z(z) = \frac{\text{Prob}\{Z \in z \pm dz\}}{\text{bivariate PDF}}
\]

\[
Y \sim f_Y(y) = \frac{\text{Prob}\{Y \in y \pm dy\}}{\text{bivariate PDF}}
\]

Bivariate density histogram = estimate of bivariate PDF
Bivariate Distributions

Bivariate CDF: $F_{ZY}(z, y) = \text{Prob}\{Z \leq z, Y \leq y\} \in [0, 1]$

Bivariate PDF: $f_{ZY}(z, y) = \frac{\partial^2 F_{ZY}(z, y)}{\partial z \partial y}$
with $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{ZY}(z, y) dz dy = 1$

Marginal distributions:
$Z : F_Z(z) = \text{Prob}\{Z \leq z\}$
$= \text{Prob}\{Z \leq z, Y \in [-\infty, +\infty]\}$
$= F_{ZY}(z, +\infty)$

$f_Z(z) = \int_{-\infty}^{+\infty} f_{ZY}(z, y) dy$

$Y : F_Y(y) = \text{Prob}\{Y \leq y\} = F_{ZY}(+\infty, y)$

Marginal moments:
$\mu_Z = E\{Z\} = \int_{-\infty}^{+\infty} z f_Z(z) dz$

$\sigma_Z = E\{(Z - \mu_Z)^2\} = \int_{-\infty}^{+\infty} (z - \mu_Z)^2 f_Z(z) dz$

Statistics of Indicators (1)

$n = 10$ joint realizations of two continuous RVs $Z$ and $Y$:

<table>
<thead>
<tr>
<th>z-values</th>
<th>2</th>
<th>1</th>
<th>5</th>
<th>1</th>
<th>6</th>
<th>3</th>
<th>9</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>y-values</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

Indicator random variables:
$I_Z(z_0) = \begin{cases} 0 & \text{if } Z > z_0 \\ 1 & \text{if } Z \leq z_0 \end{cases}$
$I_Y(y_0) = \begin{cases} 0 & \text{if } Y > y_0 \\ 1 & \text{if } Y \leq y_0 \end{cases}$

Example ($z_0 = 5, y_0 = 5$):

<table>
<thead>
<tr>
<th>z-indicators</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>y-indicators</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Univariate indicator moment (mean):

$E\{I_Z(z_0)\} = \frac{1}{N} \sum_{n=1}^{N} i_n(z_0) = F_Z(z_0) \Rightarrow F_Z(z_0 = 5) = \frac{6}{10}$

average of indicators $i_n(z_0) =$ proportion of values no greater than $z_0 = \text{CDF } F_Z(z_0)$
Statistics of Indicators (2)

$n = 10$ joint realizations of two RVs $Z$ and $Y$:

<table>
<thead>
<tr>
<th>z-values</th>
<th>2</th>
<th>1</th>
<th>5</th>
<th>1</th>
<th>6</th>
<th>3</th>
<th>9</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>y-values</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

Indicator values (example for $z_0 = 5, y_0 = 5$):

<table>
<thead>
<tr>
<th>z-indicators</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>y-indicators</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Bivariate indicator moment (mean):

$$E\{I_Z(z_0) \cdot I_Y(y_0)\} = \frac{1}{N} \sum_{n=1}^{N} i_n(z_0) \cdot i_n(y_0) = F_{ZY}(z_0, y_0)$$

Example:

$$F_{ZY}(z_0 = 5, y_0 = 5) = \frac{5}{10}$$

average of product of indicators $i_n(z_0) \cdot i_n(y_0)$ = proportion of values jointly no greater than $z_0$ and $y_0$

= bivariate CDF $F_{ZY}(z_0, y_0)$

Bivariate Moments

Covariance $\sigma_{ZY}$, a joint moment of order 2:

$$Cov\{Z, Y\} = E\{(Z - \mu_Z)(Y - \mu_Y)\} = E\{ZY\} - \mu_Z \mu_Y$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (z - \mu_Z)(y - \mu_Y) f_{ZY}(z,y) \, dz \, dy$$

Correlation:

$$\rho_{ZY} = \frac{Cov\{Z, Y\}}{\sqrt{Var\{Z\} \cdot Var\{Y\}}} \in [-1, +1]$$

- correlation = measure of linear dependence; proportional to slope of regression line: $Z = a + bY$
- $|\rho_{ZY}| = 1 \Leftrightarrow Z$ linear combination of $Y$
- independence entails lack of correlation $F_{ZY}(z,y) = F_Z(z) \cdot F_Y(y), \forall z,y \Rightarrow \rho_{ZY} = 0$
- lack of correlation does not entail independence $\rho_{ZY} = 0$ does not entail $F_{ZY}(z,y) = F_Z(z) \cdot F_Y(y)$

Moment of inertia:

$$\gamma_{ZY} = \frac{1}{2} E\{|Z - Y|^2\} = \frac{1}{2} \int \int (z - y)^2 f_{ZY}(z,y) \, dz \, dy$$

Points to remember: (i) moments summarize distributions and condense scatter plots, (ii) two RVs with same CDFs/PDFs = two RVs with same moments (reverse not always true)
Conditional Probability (1)

Conditional PDF:
\[ f_{Z|Y}(z|y) = \frac{\text{Prob}\{Z = z, Y = y\}}{\text{Prob}\{Y = y\}} = \frac{f_{ZY}(z, y)}{f_Y(y)} \]

Conditional CDF:
\[ F_{Z|Y}(z|y) = \text{Prob}\{Z \leq z|Y = y\} = \frac{\text{Prob}\{Z \leq z, Y = y\}}{\text{Prob}\{Y = y\}} \]

More generally (for two events A and B):
\[ \text{Prob}\{A|B\} = \frac{\text{Prob}\{A \text{ and } B\}}{\text{Prob}\{B\}} = \frac{\text{joint probability}}{\text{prior probability of conditioning event}} \]

Bayes’ relation:
\[ \text{Prob}\{A|B\} = \frac{\text{Prob}\{A, B\}}{\text{Prob}\{B\}} = \frac{\text{Prob}\{B|A\} \cdot \text{Prob}\{A\}}{\text{Prob}\{B\}} \]

Requisite: Probability of rain today, given rain yesterday = ?

Conditional probability:
\[ \frac{\text{Prob}\{I_n(z_n) = 1|I_{n-1}(z_{n-1}) = 1\}}{p_{1|1}} = \frac{\text{Prob}\{I_n(z_n) = 1, I_{n-1}(z_{n-1}) = 1\}}{\text{Prob}\{I_{n-1}(z_{n-1}) = 1\}} = \frac{p_{1,1}}{p_1} \]

Application: on “eroded” sequence \( \{i_n(z_n), n = 2, \ldots, 11\} \), for which both today’s and yesterday’s events available:
- \( p_{1,1} = E\{I_n(z_n) \cdot I_{n-1}(z_{n-1})\} = \frac{5}{10} \)
- \( p_1 = E\{I_{n-1}(z_{n-1})\} = E\{I_n(z_n)\} = \frac{6}{10} \)
- \( p_{1|1} = \frac{5}{6} = 0.83 \)
  Note that \( p_{1|1} = 0.83 > 0.6 = \text{proportion of rainy days} \)
  and \( p_{1,1} = 0.5 > 0.36 = p_1 \cdot p_1 \) (under independence)
**Conditional Moments**

**Conditional mean:**
\[
\mu_{Z|y} = E\{Z|Y=y\} = \int_{-\infty}^{+\infty} z \cdot f_{Z|Y}(z|y) \, dz
\]
non-linear function of \(y\)-data, in general

**Conditional variance:**
\[
\sigma_{Z|y} = Var\{Z|Y=y\} = \int_{-\infty}^{+\infty} [z - E\{Z|Y=y\}]^2 \cdot f_{Z|Y}(z|y) \, dz
\]
non-linear function of \(y\)-data, in general

**For two independent RVs:**
\[
f_{Z|Y}(z|y) = f_Z(z) \Leftrightarrow f_{Z,Y}(z,y) = f_Z(z) \cdot f_Y(y)
\]
\[
E\{Z|Y=y\} = m_Z, \quad \text{and} \quad Var\{Z|Y=y\} = \sigma_Z \text{ for any } y
\]

*under independence, if \(Y = y\), the best estimate of \(Z\) is \(m_Z\)
  (best in the least squared error sense)*

**Standard Normal Bivariate Distribution**

\[ (Z,Y) \sim N \left\{ \begin{array}{cc} 0,1 \\ 0,1 \end{array} \right\} \rho_{ZY} = \text{one single parameter} \]

**Bivariate PDF:**
\[
f_{Z,Y}(z,y) = \frac{1}{2\pi\sqrt{1-\rho_{ZY}^2}} e^{\left[ -\frac{1}{2(1-\rho_{ZY}^2)} (z^2 - 2\rho_{ZY}zy + y^2) \right]}
\]

**Marginal distributions:**
\[
Z \sim N \left\{ \begin{array}{cc} 0 = \mu_Z \\ 1 = s_Z \end{array} \right\} \quad \text{and} \quad Y \sim N \left\{ \begin{array}{cc} 0 = \mu_Y \\ 1 = s_Y \end{array} \right\}
\]

*Zero correlation entails independence:*
\[
\rho_{ZY} = 0 \Rightarrow f_{Z,Y}(z,y) = f_Z(z)f_Y(y)
\]
only for bivariate standard Normal RVs
### Bivariate Normal Distribution

\[(Z, Y) \sim N \left( \begin{array}{c} \mu_Z, \sigma_Z \\ \rho_{ZY} \\ \mu_Y, \sigma_Y \end{array} \right) \Rightarrow \left( \frac{Z - \mu_Z}{\sqrt{\sigma_Z}}, \frac{Y - \mu_Y}{\sqrt{\sigma_Y}} \right) \sim N \left( \begin{array}{c} 0, 1 \\ \rho_{ZY} \\ 0, 1 \end{array} \right) \]

Marginal distributions:

\[Z \sim N \left( \frac{\mu_Z}{\sigma_Z} \right) \quad Y \sim N \left( \frac{\mu_Y}{\sigma_Y} \right)\]

Conditional distributions (Gaussian with two parameters):

- conditional mean: \(E\{Z|Y = y\} = \mu_Z + \rho_{ZY} \frac{\sigma_Z}{\sigma_Y} (y - \mu_Y)\)
- conditional variance: \(Var\{Z|Y = y\} = \sigma_Z(1 - \rho_{ZY}^2)\)
  
conditional mean = linear function of \(y\)-data,
conditional variance = independent of \(y\)-data (homoscedastic)

### Zero correlation entails independence:

\(\rho_{ZY} = 0 \Rightarrow f_{Z|Y}(z, y) = f_Z(z)f_Y(y)\), only for bivariate Normal RVs

under independence: cond. mean = \(m_Z\), and cond. variance = \(\sigma_Z\)
under perfect correlation: cond. mean = \(\mu_Z + \frac{\sigma_Z}{\sqrt{\sigma_Y}} (y - \mu_Y)\), and
cond. variance = 0

### Multivariate Distributions (1)

\(K\) random variables: \(\{Z_1, \ldots, Z_K\}\) or a random vector \(z\)

Joint CDF:

\[F_{Z_1, \ldots, Z_K}(z_1, \ldots, z_K) = \text{Prob} \{Z_1 \leq z_1, \ldots, Z_K \leq z_K\}\]

Joint PDF:

\[f_{Z_1, \ldots, Z_K}(z_1, \ldots, z_K) = \frac{\partial^K F_{Z_1, \ldots, Z_K}(z_1, \ldots, z_K)}{\partial z_1, \ldots, \partial z_K}\]

\(K\)-th order partial derivative of joint CDF

Marginal univariate CDF:

\[F_{Z_k}(z_k) = \text{Prob} \{Z_k \leq z_k; Z_i \leq +\infty, \forall i \neq k\}\]

Marginal bivariate CDF:

\[F_{Z_kZ_{k'}}(z_k, z_{k'}) = \text{Prob} \{Z_k \leq z_k, Z_{k'} \leq z_{k'}; Z_i \leq +\infty, \forall i \neq k, k'\}\]
Multivariate Distributions (2)

$K$ random variables: $\{Z_1, \ldots, Z_K\}$

**Conditional univariate CDF:**

$$F_{Z_k|Z_i, \forall i \neq k}(z_k|Z_i = z_i, \forall i \neq k) = \text{Prob}\{Z_k \leq z_k|Z_i = z_i, \forall i \neq k\}$$

**Conditional mean:**

$$E\{Z_k|Z_i = z_i, \forall i \neq k\} = \varphi(z_i, \forall i \neq k)$$

non-linear function $\varphi(\cdot)$, in general, of $K-1$ conditioning values $\{z_i, \forall i \neq k\}$

**Independence:**

- $\text{Prob}\{Z_k \leq z_k|Z_i = z_i, \forall i \neq k\} = \text{Prob}\{Z_k \leq z_k\}, \forall k$
- $f_{Z_1, \ldots, Z_K}(z_1, \ldots, z_K) = \prod_{k=1}^{K} f_{Z_k}(z_k)$
- $f_{Z_k Z_{k'}}(z_k, z_{k'}) = f_{Z_k}(z_k) \cdot f_{Z_{k'}}(z_{k'}), \forall k, k' \Rightarrow \rho_{kk'} = 0$

pairwise independence entails lack of correlation
higher-order (e.g., 3 by 3) independence entails pairwise independence
(reverse not always true)

Multivariate Gaussian Distribution

$(K \times 1)$ random vector $z = [Z_1, \ldots, Z_K]^T$

**Univariate and bivariate moments:**

- $(K \times 1)$ mean vector: $\mu = [\mu_1, \ldots, \mu_K]^T$
- $(K \times K)$ covariance matrix: $\Sigma = E\{|z - \mu|[Z - \mu]^T\}$

**Multivariate PDF:**

$$f_z(z) = \frac{1}{(2\pi)^{K/2} |\Sigma|^{1/2}} \exp\left[ -\frac{1}{2} [z - \mu]^T \Sigma^{-1} [z - \mu] \right]$$

$$f_z(z) = f_{Z_1, \ldots, Z_K}(z_1, \ldots, z_K) = c \exp[-\frac{1}{2} Q(z_1, \ldots, z_K)],$$

$|\Sigma| = \text{determinant of } \Sigma, c = \text{constant, } Q = \text{2nd-order polynomial}$

**Properties:**

- all $(K-k')$-variate marginal distributions are Gaussian
  $\forall k' = 1, \ldots, K-1$
- all $(K-k')$-variate conditional distributions are Gaussian
- any univariate cond. distribution has (cond.) expectation which is a linear function of $z$-values, and (cond.) variance independent of $z$-values
- any linear transformation of multivariate Gaussian RVs yields multivariate Gaussian RVs
Bivariate Gaussian Distribution

Standardized versus non-standardized variables:

\[(Z, Y) \sim N \left\{ \begin{array}{l}
\mu_Z, \sigma_Z \\
\rho_{ZY} \\
\mu_Y, \sigma_Y
\end{array} \right\} \Rightarrow \left\{ \begin{array}{l}
\frac{Z - \mu_Z}{\sqrt{\sigma_Z}}, \frac{Y - \mu_Y}{\sqrt{\sigma_Y}} \\
\rho_{ZY} \\
0, 1
\end{array} \right\} \sim N \left\{ \begin{array}{l}
0, 1 \\
\rho_{ZY} \\
0, 1
\end{array} \right\}
\]

Bivariate Gaussian PDF (non-standardized):

\[
f_{ZY}(z, y) = \frac{1}{2\pi \sqrt{\sigma_Z\sigma_Y(1 - \rho_{ZY}^2)}} \exp\left[-\frac{1}{2(1 - \rho_{ZY}^2)} \left(\frac{z - \mu_Z}{\sqrt{\sigma_Z}}\right)^2 - 2\rho_{ZY} \left(\frac{z - \mu_Z}{\sqrt{\sigma_Z}}\right) \left(\frac{y - \mu_Y}{\sqrt{\sigma_Y}}\right) + \left(\frac{y - \mu_Y}{\sqrt{\sigma_Y}}\right)^2\right]
\]

a scaled and shifted version
of the standardized bivariate Gaussian PDF