Joint distributions: common distributions

Chrysafis Vogiatzis

Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

Lecture 13

ILLINOIS
ISE | Industrial & Enterprise Systems Engineering
GRAINGER COLLEGE OF ENGINEERING

Function of a random variable

Sometimes, we are interested in a random variable that is defined as a *function* of another random variable. For example:

- $Y = h(X)$, where Y is the heat of a circuit and X its current.
- $Y = h(T)$, where Y is the quality of the crop and T is the average temperature of the region.
- $Y = h(S)$, where Y is the exam grade and S is the amount of sleep the student got during the night before the exam.

Formally, let $Y = h(X)$ be a **one-to-one** transformation of a random variable X to a random variable Y.

- This means that $y = h(x)$ has a unique solution.
- Let that solution be $x = h^{-1}(y) = u(y)$.

ILLINOIS

Function of a random variable

Sometimes, we are interested in a random variable that is defined as a *function* of another random variable. For example:

- $Y = h(X)$, where Y is the heat of a circuit and X its current.
- $Y = h(T)$, where Y is the quality of the crop and T is the average temperature of the region.
- $Y = h(S)$, where Y is the exam grade and S is the amount of sleep the student got during the night before the exam.

Formally, let $Y = h(X)$ be a **one-to-one** transformation of a random variable X to a random variable Y.

- This means that $y = h(x)$ has a unique solution.
- Let that solution be $x = h^{-1}(y) = u(y)$.

Function of a random variable

Sometimes, we are interested in a random variable that is defined as a *function* of another random variable. For example:

- $Y = h(X)$, where Y is the heat of a circuit and X its current.
- $Y = h(T)$, where Y is the quality of the crop and T is the average temperature of the region.
- $Y = h(S)$, where Y is the exam grade and S is the amount of sleep the student got during the night before the exam.

Formally, let $Y = h(X)$ be a **one-to-one** transformation of a random variable X to a random variable Y.

- This means that $y = h(x)$ has a unique solution.
- Let that solution be $x = h^{-1}(y) = u(y)$.

Function of a random variable

Sometimes, we are interested in a random variable that is defined as a *function* of another random variable. For example:

- $Y = h(X)$, where Y is the heat of a circuit and X its current.
- $Y = h(T)$, where Y is the quality of the crop and T is the average temperature of the region.
- $Y = h(S)$, where Y is the exam grade and S is the amount of sleep the student got during the night before the exam.

Formally, let $Y = h(X)$ be a **one-to-one** transformation of a random variable X to a random variable Y.

- This means that $y = h(x)$ has a unique solution.
- Let that solution be $x = h^{-1}(y) = u(y)$.

Function of a random variable

Let $Y = h(X)$ be a one-to-one function of random variable X to Y. X is distributed with pmf/pdf $f_X(x)$. Then, the pmf/pdf of random variable $Y = h(X)$ can be found by:

1. Discrete X: $f_Y(y) = f_X(u(y))$.
2. Continuous X: $f_Y(y) = f_X(u(y)) \cdot |u'(y)|$,
   where $u'(y)$ is the derivative of function $u(y)$.

Example

Continuous random variable X has pdf $f(x) = \frac{x}{2}$, defined over $0 \leq x \leq 2$. What is the pdf of $Y = \sqrt{X}$?

Answer: First of all, h is a one-to-one transformation and we have $x = h^{-1}(y) = u(y) = y^2$. Then,

$$f_Y(y) = f_X(u(y)) \cdot |u'(y)| = \frac{y^2}{2} \cdot 2y = y^3.$$

Finally, X is defined over $0 \leq x \leq 2$ so $Y = \sqrt{X}$ is defined over $0 \leq y \leq \sqrt{2}$.

Function of a random variable

Let $Y = h(X)$ be a one-to-one function of random variable X to Y. X is distributed with pmf/pdf $f_X(x)$. Then, the pmf/pdf of random variable $Y = h(X)$ can be found by:

1. Discrete X: $\quad f_Y(y) = f_X(u(y))$.
2. Continuous X: $f_Y(y) = f_X(u(y)) \cdot |u'(y)|$,
   where $u'(y)$ is the derivative of function $u(y)$.

Example

Continuous random variable X has pdf $f(x) = \frac{x}{2}$, defined over $0 \leq x \leq 2$. What is the pdf of $Y = \sqrt{X}$?

Answer: First of all, h is a one-to-one transformation and we have $x = h^{-1}(y) = u(y) = y^2$. Then,

$$f_Y(y) = f_X(u(y)) \cdot |u'(y)| = \frac{y^2}{2} \cdot 2y = y^3.$$

Finally, X is defined over $0 \leq x \leq 2$ so $Y = \sqrt{X}$ is defined over $0 \leq y \leq \sqrt{2}$.

Function of a random variable

Let $Y = h(X)$ be a one-to-one function of random variable X to Y. X is distributed with pmf/pdf $f_X(x)$. Then, the pmf/pdf of random variable $Y = h(X)$ can be found by:

1. Discrete X: $f_Y(y) = f_X(u(y))$.
2. Continuous X: $f_Y(y) = f_X(u(y)) \cdot |u'(y)|$, where $u'(y)$ is the derivative of function $u(y)$.

Example

Continuous random variable X has pdf $f(x) = \frac{x}{2}$, defined over $0 \le x \le 2$. What is the pdf of $Y = \sqrt{X}$?

Answer: First of all, h is a one-to-one transformation and we have $x = h^{-1}(y) = u(y) = y^2$. Then,

$$f_Y(y) = f_X(u(y)) \cdot |u'(y)| = \frac{y^2}{2} \cdot 2y = y^3.$$

Finally, X is defined over $0 \le x \le 2$ so $Y = \sqrt{X}$ is defined over $0 \le y \le \sqrt{2}$.

Function of a random variable

Let $Y = h(X)$ be a one-to-one function of random variable X to Y. X is distributed with pmf/pdf $f_X(x)$. Then, the pmf/pdf of random variable $Y = h(X)$ can be found by:

1. Discrete X: $f_Y(y) = f_X(u(y))$.
2. Continuous X: $f_Y(y) = f_X(u(y)) \cdot |u'(y)|$,
   where $u'(y)$ is the derivative of function $u(y)$.

Example

Continuous random variable X has pdf $f(x) = \frac{x}{2}$, defined over $0 \leq x \leq 2$. What is the pdf of $Y = \sqrt{X}$?

Answer: First of all, h is a one-to-one transformation and we have $x = h^{-1}(y) = u(y) = y^2$. Then,

$$f_Y(y) = f_X(u(y)) \cdot |u'(y)| = \frac{y^2}{2} \cdot 2y = y^3.$$

Finally, X is defined over $0 \leq x \leq 2$ so $Y = \sqrt{X}$ is defined over $0 \leq y \leq \sqrt{2}$.

ILLINOIS

Function of a random variable

Let $Y = h(X)$ be a one-to-one function of random variable X to Y. X is distributed with pmf/pdf $f_X(x)$. Then, the pmf/pdf of random variable $Y = h(X)$ can be found by:

1. Discrete X: $f_Y(y) = f_X(u(y))$.
2. Continuous X: $f_Y(y) = f_X(u(y)) \cdot |u'(y)|$, where $u'(y)$ is the derivative of function $u(y)$.

Example

Continuous random variable X has pdf $f(x) = \frac{x}{2}$, defined over $0 \leq x \leq 2$. What is the pdf of $Y = \sqrt{X}$?

Answer: First of all, h is a one-to-one transformation and we have $x = h^{-1}(y) = u(y) = y^2$. Then,

$$f_Y(y) = f_X(u(y)) \cdot |u'(y)| = \frac{y^2}{2} \cdot 2y = y^3.$$

Finally, X is defined over $0 \leq x \leq 2$ so $Y = \sqrt{X}$ is defined over $0 \leq y \leq \sqrt{2}$.

ILLINOIS

Function of a random variable

Let $Y = h(X)$ be a one-to-one function of random variable X to Y. X is distributed with pmf/pdf $f_X(x)$. Then, the pmf/pdf of random variable $Y = h(X)$ can be found by:

1. Discrete X: $f_Y(y) = f_X(u(y))$.
2. Continuous X: $f_Y(y) = f_X(u(y)) \cdot |u'(y)|$, where $u'(y)$ is the derivative of function $u(y)$.

Example

Continuous random variable X has pdf $f(x) = \frac{x}{2}$, defined over $0 \leq x \leq 2$. What is the pdf of $Y = \sqrt{X}$?

Answer: First of all, h is a one-to-one transformation and we have $x = h^{-1}(y) = u(y) = y^2$. Then,

$$f_Y(y) = f_X(u(y)) \cdot |u'(y)| = \frac{y^2}{2} \cdot 2y = y^3.$$

Finally, X is defined over $0 \leq x \leq 2$ so $Y = \sqrt{X}$ is defined over $0 \leq y \leq \sqrt{2}$.

Common joint distributions

Quick review:

1. $f_{XY}(x, y)$: joint pmf/pdf.
2. marginal/conditional pmf/pdf.
3. expectations/variances.
4. independence/covariance/correlation.

Recall that all these derivations extend to more than 2 random variables.

Common joint distributions:

1. Discrete: **multinomial**.
2. Continuous: **bivariate normal** distribution.

Common joint distributions

Quick review:

1. $f_{XY}(x, y)$: joint pmf/pdf.
2. marginal/conditional pmf/pdf.
3. expectations/variances.
4. independence/covariance/correlation.

Recall that all these derivations extend to more than 2 random variables.

Common joint distributions:

1. Discrete: **multinomial**.
2. Continuous: **bivariate normal** distribution.

The multinomial distribution

Back to the binomial:

- n independent tries.
- Each try results in success or failure (2) outcomes.
- p is the probability of each try resulting in a success.
- X (number of successes) is a random variable.

Extending to the multinomial:

- Still n independent tries.
- Each try results in **multiple** (k) outcomes.
- p_i the probability of seeing outcome $i = 1, \ldots, k$.
- X_i is the number of times we see the i-th outcome.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

- $\sum_{i=1}^{k} p_i = 1.$
- $\sum_{i=1}^{k} x_i = n.$

ILLINOIS

The multinomial distribution

Back to the binomial:

- n independent tries.
- Each try results in success or failure (2) outcomes.
- p is the probability of each try resulting in a success.
- X (number of successes) is a random variable.

Extending to the multinomial:

- Still n independent tries.
- Each try results in **multiple** (k) outcomes.
- p_i the probability of seeing outcome $i = 1, \ldots, k$.
- X_i is the number of times we see the i-th outcome.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

- $\sum_{i=1}^{k} p_i = 1$.
- $\sum_{i=1}^{k} x_i = n$.

ILLINOIS

The multinomial distribution

Back to the binomial:

- n independent tries.
- Each try results in success or failure (2) outcomes.
- p is the probability of each try resulting in a success.
- X (number of successes) is a random variable.

Extending to the multinomial:

- Still n independent tries.
- Each try results in **multiple** (k) outcomes.
- p_i the probability of seeing outcome $i = 1, \ldots, k$.
- X_i is the number of times we see the i-th outcome.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3, \ldots, X_k = x_k) = \frac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

- $\sum_{i=1}^{k} p_i = 1$.
- $\sum_{i=1}^{k} x_i = n$.

ILLINOIS

Multinomial distribution: marginal/conditional pmf

Let $(X_1, X_2, \ldots, X_k)$ be a multinomial distribution with probabilities $p_1, p_2, \ldots, p_k$, respectively. Then:

- The marginal distribution of X_i is a **binomial distribution**.
  - Every X_i is binomially distributed with parameters n, p_i.
  - "What is the probability that outcome i has x_i appearances?"
- The conditional distribution of $X_1, X_2, \ldots, X_{j-1}, X_{j+1}, \ldots, X_k$ given $X_j = x_j$ is a **multinomial distribution**.
  - $X_1, X_2, \ldots, X_{j-1}, X_{j+1}, \ldots, X_k$ (that is, everything except for X_j) is multinomially distributed with parameters $n - x_j$, $q_i = \frac{p_i}{\sum_{\ell=1:\ell\neq j}^{k} p_\ell}$.
  - "What is the probability that outcomes i have x_i appearances given that X_j has appeared x_j times?"

Example

Historically, vehicles stopping at a toll station are:

- *passenger vehicles (cars) with probability 75%,*
- *commercial vehicles (trucks) with probability 15%,*
- *and motorcycles with probability 10%.*

A transportation engineer selects 10 vehicles that used the toll at random. What is the probability there were

a) *6 cars, 2 trucks, and 2 motorcycles?*

$P(X_1 = 6, X_2 = 2, X_3 = 2) = \frac{10!}{6!2!2!} 0.75^6 0.15^2 0.1^2 = 0.0505 = 5.05\%.$

b) *at most 1 motorcycle?*

Marginal distribution of X_3 *is binomial with* $n = 10, p_3 = 0.1$*:*

$P(X_3 \leq 1) = P(X_3 = 0) + P(X_3 = 1) = 0.3487 + 0.3874 = 0.7361.$

c) *6 cars and 3 trucks, given that there was 1 motorcycle?*

Conditional distribution of X_1, X_2 *given* X_3 *is multinomial with* $n = 10 - x_3 = 9, q_1 = \frac{p_1}{p_1+p_2} = \frac{5}{6}, q_2 = \frac{p_2}{p_1+p_2} = \frac{1}{6}$*:*

$P(X_1 = 6, X_2 = 3 | X_3 = 1) = \frac{9!}{6!3!} \left(\frac{5}{6}\right)^6 \left(\frac{1}{6}\right)^3$*=0.1302.*

Multinomial with $n = 10, p_1 = 0.75, p_2 = 0.15, p_3 = 0.10.$

Example

Historically, vehicles stopping at a toll station are:

- *passenger vehicles (cars) with probability 75%,*
- *commercial vehicles (trucks) with probability 15%,*
- *and motorcycles with probability 10%.*

A transportation engineer selects 10 vehicles that used the toll at random. What is the probability there were

a) *6 cars, 2 trucks, and 2 motorcycles?*

$P(X_1 = 6, X_2 = 2, X_3 = 2) = \frac{10!}{6!2!2!} 0.75^6 0.15^2 0.1^2 = 0.0505 = 5.05\%.$

b) *at most 1 motorcycle?*

Marginal distribution of X_3 is binomial with $n = 10, p_3 = 0.1$:

$P(X_3 \leq 1) = P(X_3 = 0) + P(X_3 = 1) = 0.3487 + 0.3874 = 0.7361.$

c) *6 cars and 3 trucks, given that there was 1 motorcycle?*

Conditional distribution of X_1, X_2 given X_3 is multinomial with $n = 10 - X_3 = 9, q_1 = \frac{p_1}{p_1+p_2} = \frac{5}{6}, q_2 = \frac{p_2}{p_1+p_2} = \frac{1}{6}$:

$P(X_1 = 6, X_2 = 3 | X_3 = 1) = \frac{9!}{6!3!} \left(\frac{5}{6}\right)^6 \left(\frac{1}{6}\right)^3$*=0.1302.*

Multinomial with $n = 10, p_1 = 0.75, p_2 = 0.15, p_3 = 0.10$.

Example

Historically, vehicles stopping at a toll station are:

- *passenger vehicles (cars) with probability 75%,*
- *commercial vehicles (trucks) with probability 15%,*
- *and motorcycles with probability 10%.*

A transportation engineer selects 10 vehicles that used the toll at random. What is the probability there were

a) *6 cars, 2 trucks, and 2 motorcycles?*

$P(X_1 = 6, X_2 = 2, X_3 = 2) = \frac{10!}{6!2!2!} 0.75^6 0.15^2 0.1^2 = 0.0505 = 5.05\%.$

b) *at most 1 motorcycle?*

Marginal distribution of X_3 is binomial with $n = 10, p_3 = 0.1$:

$P(X_3 \leq 1) = P(X_3 = 0) + P(X_3 = 1) = 0.3487 + 0.3874 = 0.7361.$

c) *6 cars and 3 trucks, given that there was 1 motorcycle?*

Conditional distribution of X_1, X_2 given X_3 is multinomial with $n = 10 - x_3 = 9, q_1 = \frac{p_1}{p_1+p_2} = \frac{5}{6}, q_2 = \frac{p_2}{p_1+p_2} = \frac{1}{6}$:

$P(X_1 = 6, X_2 = 3 | X_3 = 1) = \frac{9!}{6!3!} \left(\frac{5}{6}\right)^6 \left(\frac{1}{6}\right)^3$*=0.1302.*

Multinomial with $n = 10, p_1 = 0.75, p_2 = 0.15, p_3 = 0.10$.

Example

Historically, vehicles stopping at a toll station are:

- *passenger vehicles (cars) with probability 75%,*
- *commercial vehicles (trucks) with probability 15%,*
- *and motorcycles with probability 10%.*

A transportation engineer selects 10 vehicles that used the toll at random. What is the probability there were

a) *6 cars, 2 trucks, and 2 motorcycles?*
$P(X_1 = 6, X_2 = 2, X_3 = 2) = \frac{10!}{6!2!2!}0.75^6 0.15^2 0.1^2 = 0.0505 = 5.05\%$.

b) *at most 1 motorcycle?*
Marginal distribution of X_3 is binomial with $n = 10, p_3 = 0.1$:
$P(X_3 \leq 1) = P(X_3 = 0) + P(X_3 = 1) = 0.3487 + 0.3874 = 0.7361$.

c) *6 cars and 3 trucks, given that there was 1 motorcycle?*
Conditional distribution of X_1, X_2 given X_3 is multinomial with $n = 10 - x_3 = 9, q_1 = \frac{p_1}{p_1+p_2} = \frac{5}{6}, q_2 = \frac{p_2}{p_1+p_2} = \frac{1}{6}$:
$P(X_1 = 6, X_2 = 3 | X_3 = 1) = \frac{9!}{6!3!}\left(\frac{5}{6}\right)^6\left(\frac{1}{6}\right)^3$*=0.1302.*

Multinomial with $n = 10, p_1 = 0.75, p_2 = 0.15, p_3 = 0.10$.

Example

Historically, vehicles stopping at a toll station are:

- *passenger vehicles (cars) with probability 75%,*
- *commercial vehicles (trucks) with probability 15%,*
- *and motorcycles with probability 10%.*

A transportation engineer selects 10 vehicles that used the toll at random. What is the probability there were

a) *6 cars, 2 trucks, and 2 motorcycles?*
$P(X_1 = 6, X_2 = 2, X_3 = 2) = \frac{10!}{6!2!2!} 0.75^6 0.15^2 0.1^2 = 0.0505 = 5.05\%.$

b) *at most 1 motorcycle?*
Marginal distribution of X_3 *is binomial with* $n = 10, p_3 = 0.1$*:*
$P(X_3 \leq 1) = P(X_3 = 0) + P(X_3 = 1) = 0.3487 + 0.3874 = 0.7361.$

c) *6 cars and 3 trucks, given that there was 1 motorcycle?*
Conditional distribution of X_1, X_2 *given* X_3 *is multinomial with* $n = 10 - x_3 = 9, q_1 = \frac{p_1}{p_1+p_2} = \frac{5}{6}, q_2 = \frac{p_2}{p_1+p_2} = \frac{1}{6}$*:*
$P(X_1 = 6, X_2 = 3 | X_3 = 1) = \frac{9!}{6!3!} \left(\frac{5}{6}\right)^6 \left(\frac{1}{6}\right)^3$*=0.1302.*

Multinomial with $n = 10, p_1 = 0.75, p_2 = 0.15, p_3 = 0.10$.

The bivariate normal distribution

- Bivariate: two random variables X, Y.
- Normal: both normally distributed.
  - mean: μ_X, μ_Y, resp.
  - variance: σ_X^2, σ_Y^2, resp.
  - *possibly* correlated with correlation ρ_{XY}.

Then, two random variables X and Y with the above parameters are jointly distributed with a bivariate random distribution if:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot e^{\frac{-z}{2\left(1-\rho_{XY}^2\right)}},$$

where $z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$

We may contrast with the simple normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma}}.$$

ILLINOIS

The bivariate normal distribution

- Bivariate: two random variables X, Y.
- Normal: both normally distributed.
  - mean: μ_X, μ_Y, resp.
  - variance: σ_X^2, σ_Y^2, resp.
  - *possibly* correlated with correlation ρ_{XY}.

Then, two random variables X and Y with the above parameters are jointly distributed with a bivariate random distribution if:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot e^{\frac{-z}{2\left(1-\rho_{XY}^2\right)}},$$

where $z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$

We may contrast with the simple normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma}}.$$

ILLINOIS

The bivariate normal distribution

- Bivariate: two random variables X, Y.
- Normal: both normally distributed.
  - mean: μ_X, μ_Y, resp.
  - variance: σ_X^2, σ_Y^2, resp.
  - *possibly* correlated with correlation ρ_{XY}.

Then, two random variables X and Y with the above parameters are jointly distributed with a bivariate random distribution if:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot e^{\frac{-z}{2\left(1-\rho_{XY}^2\right)}},$$

where $z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$

We may contrast with the simple normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma}}.$$

ILLINOIS

The bivariate normal distribution

- Bivariate: two random variables X, Y.
- Normal: both normally distributed.
  - mean: μ_X, μ_Y, resp.
  - variance: σ_X^2, σ_Y^2, resp.
  - *possibly* correlated with correlation ρ_{XY}.

Then, two random variables X and Y with the above parameters are jointly distributed with a bivariate random distribution if:

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot e^{\frac{-z}{2\left(1-\rho_{XY}^2\right)}},$$

where $z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$

We may contrast with the simple normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma}}.$$

ILLINOIS

The bivariate normal distribution

- Bivariate: two random variables X, Y.
- Normal: both normally distributed.
  - mean: μ_X, μ_Y, resp.
  - variance: σ_X^2, σ_Y^2, resp.
  - *possibly* correlated with correlation ρ_{XY}.

Then, two random variables X and Y with the above parameters are jointly distributed with a bivariate random distribution if:

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot e^{\frac{-z}{2\left(1-\rho_{XY}^2\right)}},$$

where $z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$

We may contrast with the simple normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma}}.$$

ILLINOIS

The bivariate normal distribution

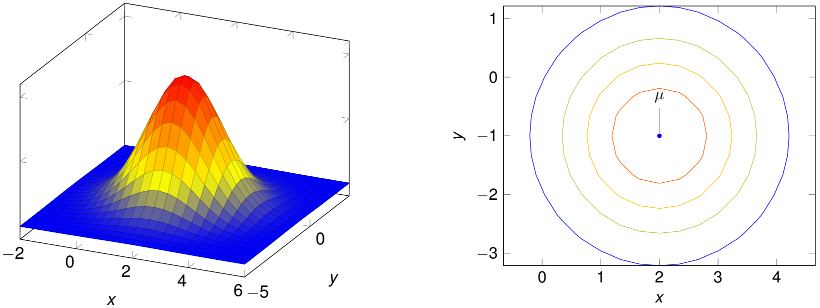
Let $\mu_X = 2, \sigma_X = 1$ and $\mu_Y = -1, \sigma_Y = 1$.

When $\rho_{XY} = 0$:

The bivariate normal distribution

Let $\mu_X = 2, \sigma_X = 1$ and $\mu_Y = -1, \sigma_Y = 1$.

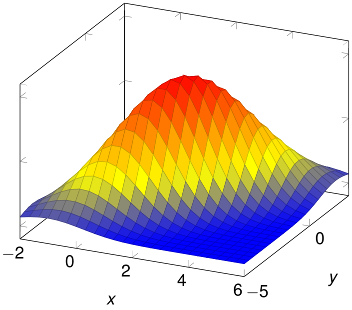
When $\rho_{XY} = 0$:

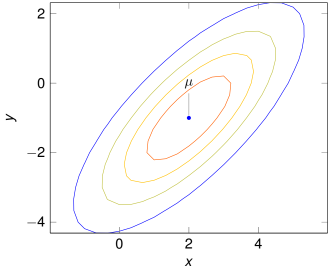


The bivariate normal distribution

Let $\mu_X = 2, \sigma_X = 1$ and $\mu_Y = -1, \sigma_Y = 1$.

When $\rho_{XY} > 0$:

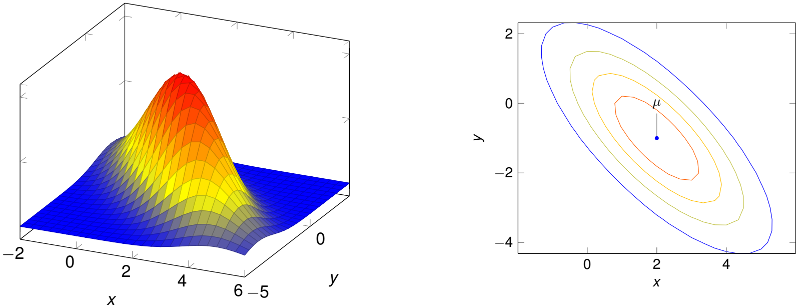




The bivariate normal distribution

Let $\mu_X = 2, \sigma_X = 1$ and $\mu_Y = -1, \sigma_Y = 1$.

When $\rho_{XY} < 0$:



Marginal and conditional distributions

The **marginal** distributions for the bivariate normal distribution are:

$$X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$$
$$Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$$

The **conditional** distribution of X given $Y = y$ is also a normal distribution with mean and variance found by:

$$\mu_{X|Y=y} = \mu_X + \rho_{XY}\left(\frac{\sigma_X}{\sigma_Y}\right)(y - \mu_Y)$$
$$\sigma^2_{X|Y=y} = \sigma_X^2\left(1 - \rho_{XY}^2\right)$$

Food for thought:

- what if X and Y are not correlated?
- what if they are *perfectly* correlated?

ILLINOIS

Marginal and conditional distributions

The **marginal** distributions for the bivariate normal distribution are:

$$X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$$
$$Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$$

The **conditional** distribution of X given $Y = y$ is also a normal distribution with mean and variance found by:

$$\mu_{X|Y=y} = \mu_X + \rho_{XY}\left(\frac{\sigma_X}{\sigma_Y}\right)(y - \mu_Y)$$
$$\sigma^2_{X|Y=y} = \sigma_X^2\left(1 - \rho_{XY}^2\right)$$

Food for thought:

- what if X and Y are not correlated?
- what if they are *perfectly* correlated?

Marginal and conditional distributions

The **marginal** distributions for the bivariate normal distribution are:

$$X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$$
$$Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$$

The **conditional** distribution of X given $Y = y$ is also a normal distribution with mean and variance found by:

$$\mu_{X|Y=y} = \mu_X + \rho_{XY}\left(\frac{\sigma_X}{\sigma_Y}\right)(y - \mu_Y)$$
$$\sigma^2_{X|Y=y} = \sigma_X^2\left(1 - \rho_{XY}^2\right)$$

Food for thought:

- what if X and Y are not correlated?
- what if they are *perfectly* correlated?

Example

Assume $X \sim \mathcal{N}(2,9)$ and $Y \sim \mathcal{N}(4,4)$ with $\rho_{XY} = 0.5$.

1. *What is $P(X \leq 1)$? What is $P(Y > 6)$?*
2. *What is $P(X \leq 1 | Y = 3)$?*

Answer:

1. X and Y follow a normal distribution. Hence:

$$P(X \leq 1) = \Phi(\frac{1-2}{3}) = \Phi(-1/3) = 0.371.$$

$$P(Y > 6) = 1 - P(Y \leq 6) = 1 - \Phi(\frac{6-4}{2}) = 1 - \Phi(1) = 0.159.$$

2. This is a conditional pdf ($X|Y = y$ is normally distributed):

$$\mu_{X|Y=y} = \mu_X + \rho_{XY}\left(\frac{\sigma_X}{\sigma_Y}\right)(y - \mu_Y) = \frac{5}{4}$$

$$\sigma^2_{X|Y=y} = \sigma^2_X\left(1 - \rho^2_{XY}\right) = \frac{27}{4}$$

Example

Assume $X \sim \mathcal{N}(2, 9)$ and $Y \sim \mathcal{N}(4, 4)$ with $\rho_{XY} = 0.5$.

1. *What is $P(X \leq 1)$? What is $P(Y > 6)$?*
2. *What is $P(X \leq 1 | Y = 3)$?*

Answer:

1. X and Y follow a normal distribution. Hence:

$$P(X \leq 1) = \Phi(\frac{1-2}{3}) = \Phi(-1/3) = 0.371.$$

$$P(Y > 6) = 1 - P(Y \leq 6) = 1 - \Phi(\frac{6-4}{2}) = 1 - \Phi(1) = 0.159.$$

2. This is a conditional pdf ($X|Y = y$ is normally distributed):

$$\mu_{X|Y=y} = \mu_X + \rho_{XY} \left(\frac{\sigma_X}{\sigma_Y}\right)(y - \mu_Y) = \frac{5}{4}$$

$$\sigma^2_{X|Y=y} = \sigma^2_X \left(1 - \rho^2_{XY}\right) = \frac{27}{4}$$

$$P(X \leq 1 | Y = 3) = \Phi\left(\frac{1 - \frac{5}{4}}{\frac{\sqrt{27}}{2}}\right) = \Phi(-0.1) = 0.46$$

ILLINOIS

Example

Assume $X \sim \mathcal{N}(2, 9)$ and $Y \sim \mathcal{N}(4, 4)$ with $\rho_{XY} = 0.5$.

1. *What is $P(X \leq 1)$? What is $P(Y > 6)$?*
2. *What is $P(X \leq 1 | Y = 3)$?*

Answer:

1. X and Y follow a normal distribution. Hence:
$$P(X \leq 1) = \Phi(\frac{1-2}{3}) = \Phi(-1/3) = 0.371.$$
$$P(Y > 6) = 1 - P(Y \leq 6) = 1 - \Phi(\frac{6-4}{2}) = 1 - \Phi(1) = 0.159.$$
2. This is a conditional pdf ($X|Y = y$ is normally distributed):
$$\mu_{X|Y=y} = \mu_X + \rho_{XY}\left(\frac{\sigma_X}{\sigma_Y}\right)(y - \mu_Y) = \frac{5}{4}$$
$$\sigma^2_{X|Y=y} = \sigma^2_X\left(1 - \rho^2_{XY}\right) = \frac{27}{4}$$
$$P(X \leq 1|Y = 3) = \Phi\left(\frac{1 - \frac{5}{4}}{\frac{\sqrt{27}}{2}}\right) = \Phi(-0.1) = 0.46$$

ILLINOIS

Example

Assume $X \sim \mathcal{N}(2,9)$ and $Y \sim \mathcal{N}(4,4)$ with $\rho_{XY} = 0.5$.

1. *What is $P(X \leq 1)$? What is $P(Y > 6)$?*
2. *What is $P(X \leq 1 | Y = 3)$?*

Answer:

1. X and Y follow a normal distribution. Hence:

$$P(X \leq 1) = \Phi(\frac{1-2}{3}) = \Phi(-1/3) = 0.371.$$

$$P(Y > 6) = 1 - P(Y \leq 6) = 1 - \Phi(\frac{6-4}{2}) = 1 - \Phi(1) = 0.159.$$

2. This is a conditional pdf ($X|Y = y$ is normally distributed):

$$\mu_{X|Y=y} = \mu_X + \rho_{XY}\left(\frac{\sigma_X}{\sigma_Y}\right)(y - \mu_Y) = \frac{5}{4}$$

$$\sigma^2_{X|Y=y} = \sigma^2_X\left(1 - \rho^2_{XY}\right) = \frac{27}{4}$$

$$P(X \leq 1 | Y = 3) = \Phi\left(\frac{1 - \frac{5}{4}}{\frac{\sqrt{27}}{2}}\right) = \Phi(-0.1) = 0.46$$

ILLINOIS

Example

Assume $X \sim \mathcal{N}(2, 9)$ and $Y \sim \mathcal{N}(4, 4)$ with $\rho_{XY} = 0.5$.

3 *What is $P(X \leq 1 \cap Y > 6)$?*

Answer: We now *have* to use the $f_{XY}(x, y)$ formula for a bivariate normal distribution. Recall that:

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot e^{\frac{-z}{2\left(1-\rho_{XY}^2\right)}},$$

where $z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$

3 Using the above:

$$P(X \leq 1 \cap Y > 6) = \int_{-\infty}^{1}\int_{6}^{\infty} f_{XY}(x, y)dydx =$$

$$= \int_{-\infty}^{1}\int_{6}^{\infty} \frac{1}{6 \cdot \sqrt{3} \cdot \pi} \cdot e^{-\frac{2}{3}\cdot\left(\frac{4x^2+9y^2-60y+8x-6xy+80}{36}\right)} =$$

$$= 0.004.$$

Example

Assume $X \sim \mathcal{N}(2,9)$ and $Y \sim \mathcal{N}(4,4)$ with $\rho_{XY} = 0.5$.

3 *What is $P(X \leq 1 \cap Y > 6)$?*

Answer: We now *have* to use the $f_{XY}(x,y)$ formula for a bivariate normal distribution. Recall that:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot e^{\frac{-z}{2\left(1-\rho_{XY}^2\right)}},$$

where $z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$

3 Using the above:

$$P(X \leq 1 \cap Y > 6) = \int_{-\infty}^{1}\int_{6}^{\infty} f_{XY}(x,y)dydx =$$

$$= \int_{-\infty}^{1}\int_{6}^{\infty} \frac{1}{6\cdot\sqrt{3}\cdot\pi} \cdot e^{-\frac{2}{3}\cdot\left(\frac{4x^2+9y^2-60y+8x-6xy+80}{36}\right)} =$$

$$= 0.004.$$

Example

Assume $X \sim \mathcal{N}(2,9)$ and $Y \sim \mathcal{N}(4,4)$ with $\rho_{XY} = 0.5$.

3 *What is $P(X \leq 1 \cap Y > 6)$?*

Answer: We now *have* to use the $f_{XY}(x,y)$ formula for a bivariate normal distribution. Recall that:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \cdot e^{\frac{-z}{2\left(1-\rho_{XY}^2\right)}},$$

where $z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$

3 Using the above:

$$P(X \leq 1 \cap Y > 6) = \int_{-\infty}^{1}\int_{6}^{\infty} f_{XY}(x,y)dydx =$$

$$= \int_{-\infty}^{1}\int_{6}^{\infty} \frac{1}{6\cdot\sqrt{3}\cdot\pi} \cdot e^{-\frac{2}{3}\cdot\left(\frac{4x^2+9y^2-60y+8x-6xy+80}{36}\right)} =$$

$$= 0.004.$$