50.034 - Introduction to Probability and Statistics

Week 5 – Lecture 10

January–May Term, 2019

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Lecture 10 Assigned Readings: 4.4, 5.6, 5.10

Outline of Lecture

- Moments
- Moment generating function
- Gaussian/Normal distribution
- Standard normal distribution
- Bivariate normal distribution

Mean, Variance, and beyond

Suppose X is an arbitrary random variable.

The mean and variance of X can be thought of as two different descriptive summaries of the probability distribution of X.

Mean and variance are NOT the only two possible descriptive summaries for probability distributions!

There is a long infinite list of possible descriptive summaries of the distribution of X called moments.

- Mean and variance are two special kinds of moments.

Moments and Central Moments

There are 2 types of moments: **moments**, and **central moments**.

Let X be any arbitrary random variable.

Definition: For every positive integer k, the expectation $\mathbf{E}[X^k]$ is called the k-th moment of X.

- In particular, this means that the mean of X is exactly the same as the first moment of X.

Definition: If X has mean μ, then for every positive integer k, the expectation $\mathbf{E}[(X-\mu)^k]$ is called the k-th central moment of X.

- **Recall:** The variance of X is $\mathrm{var}(X) = \mathbf{E}[(X-\mu)^2]$.
- In other words, the variance of X is exactly the same as the second central moment of X.
- The first central moment is exactly zero.

Note: "central moment" is also called "moment about the mean".

Existence of Moments

Recall: (From Lecture 4) For $\mathbf{E}[X]$ to exist, there is a technical condition that has to be satisfied.

- If X is a discrete R.V., then $\mathbf{E}[X]$ exists if:

$$\sum_{x \in D_{\geq 0}} x \cdot p(x) < \infty \quad \text{or} \quad \sum_{x \in D_{<0}} (-x) \cdot p(x) < \infty \quad \text{(or both)}.$$

- If X is a continuous R.V., then $\mathbf{E}[X]$ exists if:

$$\int_0^\infty x \cdot f(x)\, dx < \infty \quad \text{or} \quad \int_{-\infty}^0 (-x) \cdot f(x)\, dx < \infty \quad \text{(or both)}.$$

Similarly, for moments to exist, there are technical conditions too! For any R.V. X:

- The k-th moment of X exists if $\mathbf{E}[|X|^k] < \infty$.
- If X has mean μ, then the k-th central moment of X exists if $\mathbf{E}[|X - \mu|^k] < \infty$

Moment generating function

Definition: Let X be any random variable. For each real number t, define

$$(t) = \mathbf{E}[e^{tX}].$$

The function $\psi(t)$ is called the moment generating function (mgf) of X, provided the expectation $\mathbf{E}[e^{tX}]$ exists.

- The mgf of X (if it exists) depends only on the distribution of X.
- If X and Y are R.V.'s with the same distribution, then X and Y must have the same mgf (if it exists).

Why is $\boldsymbol{\Psi}(\mathbf{t})$ called "moment generating function"?

Theorem: Let X be a R.V. such that its mgf $\psi(t)$ exists and is finite for all values of t in some open interval around the point $t = 0$. Then for every positive integer k, the **k-th moment** of X exists, is finite, and equals the **k-th derivative** $\psi^{(k)}(0)$ at $t = 0$.

- The derivatives of $\psi(t)$ at $t = 0$ generate the moments of X.

Example 1

Let X be a discrete R.V. with pmf $p(x)$ given by

$$p(x) = \begin{cases} \frac{3^x e^{-3}}{x!}, & \text{if } x = 0, 1, 2, \ldots; \\ 0, & \text{otherwise;} \end{cases}$$

(Recall: This means that X is a Poisson R.V. with parameter 3.)

Find the mgf of X. (Hint: $\displaystyle\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$.)

Example 1 - Solution

Recall: (Lecture 5) For any discrete R.V. with pmf $p(x)$, such that D is the set of possible values, and for any function $h : \mathbb{R} \to \mathbb{R}$,

$$\mathbf{E}[h(X)] = \sum_{x \in D} h(x) \cdot p(x).$$

In this example, $D = \{0, 1, 2, \ldots\}$.

Thus, for each positive integer t,

$$\begin{aligned}
\mathbf{E}[e^{tX}] &= \sum_{x=0}^{\infty} e^{tx} \cdot p(x) = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{3^x e^{-3}}{x!} \\
&= e^{-3} \sum_{x=0}^{\infty} \frac{(3e^t)^x}{x!} = e^{-3} e^{(3e^t)} \\
&= e^{3(e^t - 1)}.
\end{aligned}$$

Therefore, the mgf of X is $\psi(t) = e^{3(e^t - 1)}$.

Moments and distributions

In Example 1 of the previous lecture, we saw that two R.V.'s could be different but still have the same mean and variance.

- In other words, the first moment and second central moment are insufficient for determining the distribution of a R.V.

However, if we know the **moment generating function**, then we can completely determine the distribution of the R.V.

Theorem: Let X and Y be R.V.'s, and suppose they have mgf's $\psi_X(t)$ and $\psi_Y(t)$ respectively. If $\psi_X(t)$ and $\psi_Y(t)$ are finite and identical for all values of t in some open interval around the point $t = 0$, then the distributions of X and Y must be identical.

- In other words, the mgf (if it exists) is another way to represent the same information given by the probability distribution of a R.V.

Note: Given the mgf, we can easily calculate mean and variance.

Example 2

Let X be a R.V. whose mgf exists and is given by

$$\psi(t) = \frac{1}{4(1 - \frac{1}{2}e^t)^2}$$

for all $-\infty < t < \infty$.

- Find the expectation of X.
- Determine the value of $\mathbf{E}[X^2]$.
- Find the variance of X.

Example 2

Solution: We are given that $\psi(t) = \dfrac{1}{4(1 - \frac{1}{2}e^t)^2}$.

First, we compute the first derivative:

$$\psi'(t) = \frac{d}{dt}\psi(t) = \frac{e^t}{4(1 - \frac{1}{2}e^t)^3}.$$

Hence $\mathbf{E}[X] = \psi'(0) = \frac{1}{4(\frac{1}{8})} = 2$.

Next, we compute the second derivative:

$$\psi''(t) = \tfrac{d}{dt}\psi'(t) = \frac{e^t(e^t + 1)}{4(1 - \frac{1}{2}e^t)^4}.$$

Hence $\mathbf{E}[X^2] = \psi''(0) = \frac{1(2)}{4(\frac{1}{16})} = 8$.

Finally, $\mathrm{var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = 8 - 2^2 = 4$.

μ, σ^2, and mgf's of several probability distributions

Bernoulli distribution (with parameter p) *[discrete R.V.]*

$$\mu = p \text{ and } \sigma^2 = p(1-p) \text{ and } \psi(t) = 1 - p + pe^t.$$

Binomial distribution (with parameters n and p) *[discrete R.V.]*

$$\mu = np \text{ and } \sigma^2 = np(1-p) \text{ and } \psi(t) = (1 - p + pe^t)^n.$$

Geometric distribution (with parameter p) *[discrete R.V.]*

$$\mu = \frac{1-p}{p} \text{ and } \sigma^2 = \frac{1-p}{p^2} \text{ and } \psi(t) = \frac{p}{1-(1-p)e^t}.$$

Poisson distribution (with parameter λ) *[discrete R.V.]*

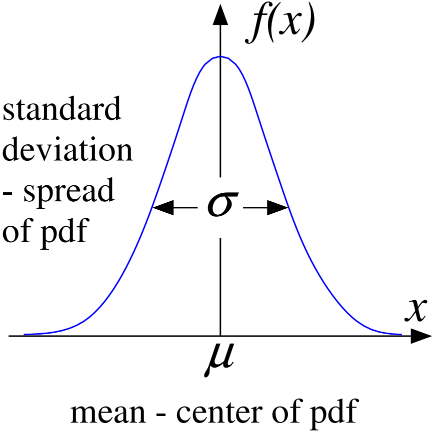
$$\mu = \lambda \text{ and } \sigma^2 = \lambda \text{ and } \psi(t) = e^{\lambda(e^t - 1)}.$$

Exponential distribution (with parameter λ) *[continuous R.V.]*

$$\mu = \frac{1}{\lambda} \text{ and } \sigma^2 = \frac{1}{\lambda^2} \text{ and } \psi(t) = \frac{\lambda}{\lambda - t} \quad (\text{for } t < \lambda).$$

Gaussian/Normal distribution

The most widely used probability distribution is called the Gaussian distribution, or also called the normal distribution.



Carl Friedrich Gauss (1777 - 1855)

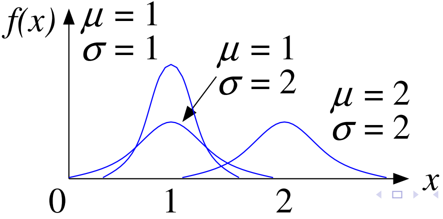
Gaussian/Normal distribution

A continuous R.V. X is called Gaussian or normal if its pdf is given by

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

for some parameters μ and σ satisfying $-\infty < \mu < \infty$ and $\sigma > 0$.

- The pdf is also often written as $f(x; \mu, \sigma)$.
- We say that X is the Gaussian R.V. with parameters μ and σ, or the normal R.V. with parameters μ and σ.
- Its distribution is called Gaussian (or normal) distribution.
- The graph of its pdf is sometimes called a "bell-shaped curve".
  - This graph is symmetric about $x = \mu$, shaped like a "bell".



Why is normal distribution called "normal"?

Reason 1:
Many real-world R.V.'s studied in various physical experiments have distributions that are approximately Gaussian/normal.

Some Examples:

- Heights of individuals in a homogeneous population of people.
- Weights of Fuji apples harvested from Fujisaki, Japan.
- Tensile strength of pieces of steel produced in a factory.

Reason 2:
We will later see in Lecture 12 the following very general fact:

- Suppose we take a **large** random sample from **any** distribution (continuous, discrete, or mixed) with finite mean and variance.
- Even if the distribution is not close to a normal distribution, it is a fascinating fact that the **sample mean** would always approximately follow a normal distribution.
  (Remember: Different samples give different sample means.)

More remarks on normal distributions

Recall: If X is a normal R.V. with parameters μ and σ, then its pdf is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad \text{for } -\infty < x < \infty.$$

- For any function $f(x)$, we write $\exp(f(x))$ to mean $e^{f(x)}$.
- **Fact:** The mean of X equals μ.
- **Fact:** The variance of X equals σ^2.

Consequently, we usually say one of the following:

- X has the normal distribution with mean μ and variance σ^2.
- X is normally distributed with mean μ and variance σ^2.

Common notation: We write $X \sim N(\mu, \sigma^2)$ to mean that X is normally distributed with mean μ and variance σ^2.

Theorem: If $X \sim N(\mu, \sigma^2)$, then its mgf is $\psi(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$.

Linear functions of normal R.V.'s

Theorem: If $X \sim N(\mu, \sigma^2)$, and if $Y = aX + b$ for some constants a, b such that $a \neq 0$, then Y has the normal distribution with mean $a\mu + b$ and variance $a^2\sigma^2$, i.e. $Y \sim N(a\mu + b, a^2\sigma^2)$.

- In other words, any linear function of a normal R.V., that is not the zero function, is always a normal R.V.!

Theorem: Suppose $X_1, \ldots, X_n$ are **independent** R.V.'s, such that $X_i \sim N(\mu_i, \sigma_i^2)$ for each $i = 1, \ldots, n$. If $a_1, \ldots, a_n, b$ are constants such that at least one of $a_1, \ldots, a_n$ is non-zero, then the R.V. $a_1X_1 + \cdots + a_nX_n + b$ has the normal distribution with mean $a_1\mu_1 + \cdots + a_n\mu_n + b$ and variance $a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2$.

- In other words, any linear function of several normal R.V.'s, that is not the zero function, is always a normal R.V.!

Very important special case:
Corollary: If $X \sim N(\mu, \sigma^2)$, then the R.V. $Z = \frac{X-\mu}{\sigma}$ has the normal distribution with mean 0 and variance 1, i.e. $Z \sim N(0, 1)$.

Special Case: Standard normal distribution

The standard normal distribution is the normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$.

- A standard normal R.V. is a R.V. with the standard normal distribution, i.e. its pdf is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad \text{for } -\infty < x < \infty.$$

Common Notation:

- A standard normal R.V. is usually denoted by Z.
  - We write $Z \sim N(0, 1)$ to mean Z is a standard normal R.V.
- The **pdf** of Z is usually written as $\phi(z)$.
  - $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}z^2), \quad \text{for } -\infty < z < \infty.$
- The **cdf** of Z is usually written as $\Phi(z)$.
  - Φ is the capitalization of the greek letter ϕ ("phi").
  - $\Phi(z)$ is usually either called the standard normal cdf or the standard normal distribution function.

Properties of standard normal distribution

Suppose Z is a standard normal R.V., i.e. $Z \sim N(0,1)$.
Let $\phi(z)$ be the pdf of Z, and let $\Phi(z)$ be the cdf of Z.

Fact: The graph of $\phi(z)$ is symmetric about the point $z = 0$.

- In other words, the pdf is symmetric about its mean 0.

Fact: $\Phi(-z) = 1 - \Phi(z)$ for all real numbers z.

- This follows from the symmetry of the pdf $\phi(z)$.
- In other words, $\Pr(Z \leq z) = \Pr(Z \geq -z)$ for all $z \in \mathbb{R}$.
- **Note:** $0 < \Phi(z) < 1$ for all real numbers z.

Fact: $\Phi^{-1}(p) = -\Phi^{-1}(1-p)$ for all real numbers $0 < p < 1$.

- To see why, let $z = \Phi^{-1}(p)$ in previous fact, and apply Φ^{-1}.

Fact: (Very useful corollary from 2 slides back) If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma}$ has the standard normal distribution.

Converting normal R.V. to standard normal R.V.

Let $X \sim N(\mu, \sigma^2)$, and let $F(x)$ be the cdf of X.
Let $Z \sim N(0, 1)$, and let $\Phi(z)$ be the cdf of Z.

We know that $\frac{X-\mu}{\sigma}$ and Z have the exact same distributions.

Theorem:

- $F(x) = \Phi(\frac{x-\mu}{\sigma})$ for all real numbers x.
- $F^{-1}(p) = \mu + \sigma\Phi^{-1}(p)$ for all real numbers $0 < p < 1$.

Important Consequence:
To calculate the cdf $F(x)$ of any normal R.V. $X \sim N(\mu, \sigma^2)$ at any given value x, we only need to know how to calculate the cdf $\Phi(z)$ of the standard normal R.V. for the value $z = \frac{x-\mu}{\sigma}$.

- Similarly, to calculate $F^{-1}(p)$, we only need to know $\Phi^{-1}(p)$.

Computing the cdf of a normal R.V.

If $X \sim N(\mu, \sigma^2)$, then by definition, its cdf is

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-\mu)^2}{2\sigma^2}}\,du.$$

- **Problem:** This integral has no closed form formula!
- Hence, the cdf $F(x)$ can only be computed approximately, using integral approximation methods (e.g. trapezoidal rule).

Workaround to problem: Use the formula $F(x) = \Phi(\frac{x-\mu}{\sigma})$.

- To use this formula, we first compute approximations to $\Phi(z)$ for various possible z, then store all these computed values. To determine $F(x)$, we just "look up" the value of $\Phi(\frac{x-\mu}{\sigma})$.
  - (Before there were computers:) Computed values are stored as tables of values. For example, there is a table at the back of the course textbook for values $\Phi(0.01), \Phi(0.02), \ldots, \Phi(4.00)$.
  - (Today:) Many calculators and statistical packages have these values stored in lookup tables. So for example, when you query $\Phi(0.05)$, the stored value for $\Phi(0.05)$ would be retrieved.

Table of stored values in back of course textbook

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{1/2}} \exp\left(-\tfrac{1}{2}u^2\right) du$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.5000	0.60	0.7257	1.20	0.8849	1.80	0.9641	2.40	0.9918
0.01	0.5040	0.61	0.7291	1.21	0.8869	1.81	0.9649	2.41	0.9920
0.02	0.5080	0.62	0.7324	1.22	0.8888	1.82	0.9656	2.42	0.9922
0.03	0.5120	0.63	0.7357	1.23	0.8907	1.83	0.9664	2.43	0.9925
0.04	0.5160	0.64	0.7389	1.24	0.8925	1.84	0.9671	2.44	0.9927
0.05	0.5199	0.65	0.7422	1.25	0.8944	1.85	0.9678	2.45	0.9929
0.06	0.5239	0.66	0.7454	1.26	0.8962	1.86	0.9686	2.46	0.9931
0.07	0.5279	0.67	0.7486	1.27	0.8980	1.87	0.9693	2.47	0.9932
0.08	0.5319	0.68	0.7517	1.28	0.8997	1.88	0.9699	2.48	0.9934
0.09	0.5359	0.69	0.7549	1.29	0.9015	1.89	0.9706	2.49	0.9936
0.10	0.5398	0.70	0.7580	1.30	0.9032	1.90	0.9713	2.50	0.9938
0.11	0.5438	0.71	0.7611	1.31	0.9049	1.91	0.9719	2.52	0.9941
0.12	0.5478	0.72	0.7642	1.32	0.9066	1.92	0.9726	2.54	0.9945
0.13	0.5517	0.73	0.7673	1.33	0.9082	1.93	0.9732	2.56	0.9948
0.14	0.5557	0.74	0.7704	1.34	0.9099	1.94	0.9738	2.58	0.9951
0.15	0.5596	0.75	0.7734	1.35	0.9115	1.95	0.9744	2.60	0.9953
0.16	0.5636	0.76	0.7764	1.36	0.9131	1.96	0.9750	2.62	0.9956
0.17	0.5675	0.77	0.7794	1.37	0.9147	1.97	0.9756	2.64	0.9959
0.18	0.5714	0.78	0.7823	1.38	0.9162	1.98	0.9761	2.66	0.9961
0.19	0.5753	0.79	0.7852	1.39	0.9177	1.99	0.9767	2.68	0.9963
0.20	0.5793	0.80	0.7881	1.40	0.9192	2.00	0.9773	2.70	0.9965
0.21	0.5832	0.81	0.7910	1.41	0.9207	2.01	0.9778	2.72	0.9967
0.22	0.5871	0.82	0.7939	1.42	0.9222	2.02	0.9783	2.74	0.9969
0.23	0.5910	0.83	0.7967	1.43	0.9236	2.03	0.9788	2.76	0.9971
0.24	0.5948	0.84	0.7995	1.44	0.9251	2.04	0.9793	2.78	0.9973
0.25	0.5987	0.85	0.8023	1.45	0.9265	2.05	0.9798	2.80	0.9974
0.26	0.6026	0.86	0.8051	1.46	0.9279	2.06	0.9803	2.82	0.9976
0.27	0.6064	0.87	0.8079	1.47	0.9292	2.07	0.9808	2.84	0.9977
0.28	0.6103	0.88	0.8106	1.48	0.9306	2.08	0.9812	2.86	0.9979
0.29	0.6141	0.89	0.8133	1.49	0.9319	2.09	0.9817	2.88	0.9980

Another table with a different format

$\Phi(1.25) = 0.89435.$

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189

Example 3

Suppose $Z \sim N(0,1)$, i.e. Z has the standard normal distribution. Determine the value of $\Pr(-0.71 < Z < 1.26)$.

Solution:
Let $\Phi(z)$ be the standard normal cdf. Note that

$$\begin{aligned}
\Pr(-0.71 < Z < 1.26) &= \Pr(-0.71 < Z \leq 1.26) \\
&= \Pr(Z \leq 1.26) - \Pr(Z \leq -0.71) \\
&= \Phi(1.26) - \Phi(-0.71) \\
&= \Phi(1.26) - (1 - \Phi(0.71)) \\
&= \Phi(1.26) + \Phi(0.71) - 1.
\end{aligned}$$

Here, we used the fact that $\Phi(-z) = 1 - \Phi(z)$ for all $z \in \mathbb{R}$.

Example 3

From the table, $\Phi(1.26) \approx 0.8962$, and $\Phi(0.71) \approx 0.7611$. Therefore,

$$\begin{aligned}\Pr(-0.71 < Z < 1.26) &= \Phi(1.26) + \Phi(0.71) - 1 \\ &\approx 0.8962 + 0.7611 - 1 \\ &= 0.6573.\end{aligned}$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{1/2}} \exp\left(-\tfrac{1}{2}u^2\right) du$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.5000	0.60	0.7257	1.20	0.8849	1.80	0.9641	2.40	0.9918
0.01	0.5040	0.61	0.7291	1.21	0.8869	1.81	0.9649	2.41	0.9920
0.02	0.5080	0.62	0.7324	1.22	0.8888	1.82	0.9656	2.42	0.9922
0.03	0.5120	0.63	0.7357	1.23	0.8907	1.83	0.9664	2.43	0.9925
0.04	0.5160	0.64	0.7389	1.24	0.8925	1.84	0.9671	2.44	0.9927
0.05	0.5199	0.65	0.7422	1.25	0.8944	1.85	0.9678	2.45	0.9929
0.06	0.5239	0.66	0.7454	1.26	0.8962	1.86	0.9686	2.46	0.9931
0.07	0.5279	0.67	0.7486	1.27	0.8980	1.87	0.9693	2.47	0.9932
0.08	0.5319	0.68	0.7517	1.28	0.8997	1.88	0.9699	2.48	0.9934
0.09	0.5359	0.69	0.7549	1.29	0.9015	1.89	0.9706	2.49	0.9936
0.10	0.5398	0.70	0.7580	1.30	0.9032	1.90	0.9713	2.50	0.9938
0.11	0.5438	0.71	0.7611	1.31	0.9049	1.91	0.9719	2.52	0.9941
0.12	0.5478	0.72	0.7642	1.32	0.9066	1.92	0.9726	2.54	0.9945
0.13	0.5517	0.73	0.7673	1.33	0.9082	1.93	0.9732	2.56	0.9948
0.14	0.5557	0.74	0.7704	1.34	0.9099	1.94	0.9738	2.58	0.9951
0.15	0.5596	0.75	0.7734	1.35	0.9115	1.95	0.9744	2.60	0.9953

Example 4

Let X be R.V. that has the normal distribution with mean 1 and variance 25. Determine the value of $\Pr(-2.25 < X < 10)$.

Solution:
First, we notice that

$$\begin{aligned}
\Pr(-2.25 < X < 10) &= \Pr\left(\frac{-2.25-1}{5} < \frac{X-1}{5} < \frac{10-1}{5}\right) \\
&= \Pr\left(-0.65 < \frac{X-1}{5} < 1.8\right) \\
&= \Pr\left(-0.65 < \frac{X-1}{5} \leq 1.8\right)
\end{aligned}$$

Example 4

Let $\Phi(z)$ be the standard normal cdf, and define the R.V. $Z = \frac{X-1}{5}$.
Since $X \sim N(1, 25)$, it follows that $Z \sim N(0, 1)$ i.e. Z has the standard normal distribution.

Thus, we get

$$\begin{aligned}
\Pr(-2.25 < X < 10) &= \Pr(-0.65 < Z \leq 1.8) \\
&= \Pr(Z \leq 1.8) - \Pr(Z \leq -0.65) \\
&= \Phi(1.8) - \Phi(-0.65) \\
&= \Phi(1.8) - (1 - \Phi(0.65)) \\
&= \Phi(1.8) + \Phi(0.65) - 1.
\end{aligned}$$

Example 4

From the table, $\Phi(1.80) \approx 0.9641$, and $\Phi(0.65) \approx 0.7422$. Therefore,

$$\begin{aligned}
\Pr(-2.25 < X < 10) &= \Phi(1.8) + \Phi(0.65) - 1 \\
&\approx 0.9641 + 0.7422 - 1 \\
&= 0.7063.
\end{aligned}$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{1/2}} \exp\left(-\tfrac{1}{2}u^2\right) du$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.5000	0.60	0.7257	1.20	0.8849	1.80	0.9641	2.40	0.9918
0.01	0.5040	0.61	0.7291	1.21	0.8869	1.81	0.9649	2.41	0.9920
0.02	0.5080	0.62	0.7324	1.22	0.8888	1.82	0.9656	2.42	0.9922
0.03	0.5120	0.63	0.7357	1.23	0.8907	1.83	0.9664	2.43	0.9925
0.04	0.5160	0.64	0.7389	1.24	0.8925	1.84	0.9671	2.44	0.9927
0.05	0.5199	0.65	0.7422	1.25	0.8944	1.85	0.9678	2.45	0.9929
0.06	0.5239	0.66	0.7454	1.26	0.8962	1.86	0.9686	2.46	0.9931
0.07	0.5279	0.67	0.7486	1.27	0.8980	1.87	0.9693	2.47	0.9932
0.08	0.5319	0.68	0.7517	1.28	0.8997	1.88	0.9699	2.48	0.9934
0.09	0.5359	0.69	0.7549	1.29	0.9015	1.89	0.9706	2.49	0.9936
0.10	0.5398	0.70	0.7580	1.30	0.9032	1.90	0.9713	2.50	0.9938
0.11	0.5438	0.71	0.7611	1.31	0.9049	1.91	0.9719	2.52	0.9941
0.12	0.5478	0.72	0.7642	1.32	0.9066	1.92	0.9726	2.54	0.9945
0.13	0.5517	0.73	0.7673	1.33	0.9082	1.93	0.9732	2.56	0.9948
0.14	0.5557	0.74	0.7704	1.34	0.9099	1.94	0.9738	2.58	0.9951
0.15	0.5596	0.75	0.7734	1.35	0.9115	1.95	0.9744	2.60	0.9953

From normal R.V. to standard normal R.V.

Let $X \sim N(\mu, \sigma^2)$, let $Z \sim N(0,1)$, and suppose Z has cdf $\Phi(z)$. As the previous example shows, we can write $\Pr(a \le X \le b)$ in terms of the cdf of Z:

$$\begin{aligned}\Pr(a \le X \le b) &= \Pr\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\end{aligned}$$

- This is because $a \le x \le b$ if and only if $\frac{a-\mu}{\sigma} \le \frac{x-\mu}{\sigma} \le \frac{b-\mu}{\sigma}$.
  - Any number x that satisfies $a \le x \le b$ must also satisfy $\frac{a-\mu}{\sigma} \le \frac{x-\mu}{\sigma} \le \frac{b-\mu}{\sigma}$, and vice versa.

We also have

$$\Pr(X \ge a) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$\Pr(X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

Intuition for bivariate normal distributions

Many real-world R.V.'s can be modeled by the normal distribution. Such R.V.'s may not be independent.

- For example, the height and weight of individuals have distributions that are each approximately normal.
- Let X and Y be normal R.V.'s representing the height and weight respectively.
- Studies have shown that X and Y are not independent.

To better understand these two normal R.V.'s depend on each other, we need to consider their **joint distribution**.

In general, to understand the joint distribution of two normal R.V.'s, we need to look at its bivariate normal distribution.

Bivariate normal distributions

Let X and Y be continuous R.V.'s.
Let $\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho$ be real constants such that $-\infty < \mu_X < \infty$, $-\infty < \mu_Y < \infty$, and $\sigma_X > 0$, $\sigma_Y > 0$, and $-1 < \rho < 1$.

If the joint pdf $f(x, y)$ of X and Y equals

$$\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right),$$

then we say that X and Y have a bivariate normal distribution.

Some Facts: (See course textbook for proofs)

- Given this joint pdf $f(x, y)$, X and Y must be normal R.V.'s.
  - $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$.
- The correlation of X and Y must be ρ.
  - **Recall:** Correlation is $\rho(X, Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{\mathbf{E}[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\sigma_Y}$.

So, we say more precisely that X and Y have the bivariate normal distribution with means μ_X, μ_Y, variances σ_X^2, σ_Y^2, and correlation ρ.

Independence and bivariate normal distributions

Recall: (Lecture 9) For two R.V.'s, independence implies zero covariance, but zero covariance does not imply independence.

- Example 2 of Lecture 9 gives two R.V.'s that have zero covariance but are dependent.
- **Consequence:** In general, two R.V.'s with zero correlation are not necessarily independent.

However, if X and Y are continuous R.V.'s with a **bivariate normal distribution**, then we have the following nice theorem.

Theorem: Let X and Y be continuous R.V.'s with a bivariate normal distribution. Then X and Y are independent if and only if they have zero correlation.

- In other words, for the special case of bivariate normal distributions, zero correlation does imply independence!
- So in this case, we can use $\rho(X, Y)$ to check for independence.

Summary

- Moments
- Moment generating function
- Gaussian/Normal distribution
- Standard normal distribution
- Bivariate normal distribution

Reminder:
There is **mini-quiz 2** (15mins) next week during Cohort Class.

- Tested on all materials from Lectures 6–10 and Cohort classes weeks 3–5.
- Today's lecture is Lecture 10.