STA 103 Lecture 6: Continuous Random Variables

Instructor: Wookyeong Song

Department of Statistics, University of California, Davis

Aug 18th, 2025

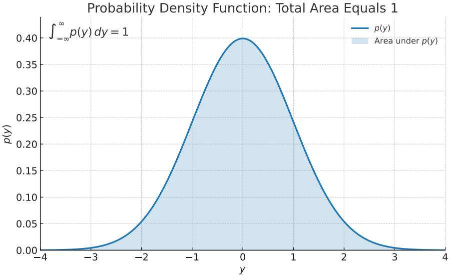
Continuous Random Variables

- Until now all the RVs we have worked with are discrete, i.e., you can list all possible values.
- In this case, we have seen that the probability mass function (PMF) characterizes the random variable.
- Discrete random variables X are characterized by a PMF $p(x)$ in a either table or function form.
- When the possible values make a continuous of numbers, we say the random variable is continuous.

Continuous Random Variables

- Continuous random variables Y are characterized by a probability density function (PDF) $p(y)$.

- Specifically, the PDF must satisfy the following properties:
  - $p(y) \geq 0$, for all y, and
  - $\int_{-\infty}^{\infty} p(y) dy = 1$ (the total area under the curve is 1).

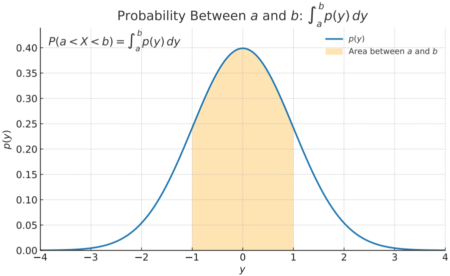


Continuous Random Variables

- The probability that the random variable X falls between two values a and b is given by the area under the density curve between those two points:

$$P(a < Y < b) = \int_a^b p(y)dy.$$

- For example, when $a = -1$, and $b = 1$, we have the following:



PDF Property: Zero Probability at Single Point

- For a continuous random variable Y with PDF $p(y)$, the probability that Y takes on any specific value c is

$$P(Y = c) = \int_c^c p(y)dy = 0.$$

- This may seem unintuitive at first, but it is a natural consequence of how continuous distribution works.

- Assigning a positive probability to any individual value would lead to an infinite total probability over an interval, which is **not** valid.

PDF Property: Zero Probability at Single Point

- As a result, for continuous random variables:

$$P(a < Y < b) = P(a \leq Y < b) = P(a < Y \leq b) = P(a \leq Y \leq b).$$

- All **four** forms are equivalent because the probability at a single point like $(Y = a)$ or $(Y = b)$ is zero.

- Note: this is **not** true for discrete random variables, where individual values can have positive probability.

The main difference between PMF and PDF

- The main difference is how one computes probabilities and expectation.

- For discrete X, and some possible outcome set A:

$$P(X \in A) = \sum_{x \in A} p(x),$$
$$E(X) = \sum_{x} x p(x),$$
$$E(f(X)) = \sum_{x} f(x) p(x).$$

- For continuous Y, and some possible outcome set B:

$$P(Y \in B) = \int_{B} p(y) dy,$$
$$E(Y) = \int_{-\infty}^{\infty} y p(y) dy,$$
$$E(f(Y)) = \int_{-\infty}^{\infty} f(y) p(y) dy.$$

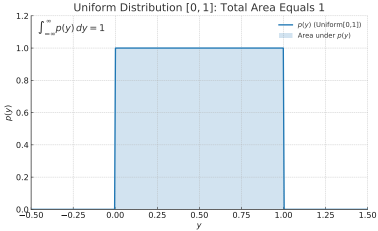
Uniform Random Variable

- **Definition**: A **uniform random variable** is used to model situations where all outcomes in a certain interval are equally likely.

- For example, choosing a number at random between 0 and 1 implies a uniform distribution on the interval $[0, 1]$, where the probability of any subinterval of length $0 < l < 1$ is exactly l.

- The probability density function (PDF) for a uniform distribution on [0,1] is:

$$f(y) = \begin{cases} 1, & 0 \leq y \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

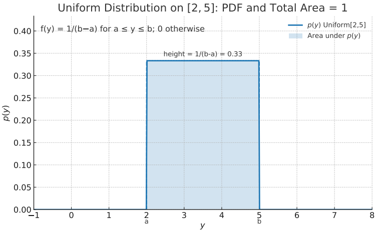


Uniform Random Variable

- More generally, the uniform distribution on an interval $[a, b]$ has the PDF:

$$f(y) = \begin{cases} \frac{1}{(b-a)}, & a \leq y \leq b, \\ 0, & \text{otherwise.} \end{cases}$$

- For example, when $a = 2$, and $b = 5$, we have the following:



- This implies that any subinterval within $[a, b]$ has a probability proportional to its length.

Mean and Variance of Uniform Distribution

- Let $Y \sim \text{Uniform}(a, b)$ with probability density function:

$$f(y) = \begin{cases} \frac{1}{(b-a)}, & a \leq y \leq b, \\ 0, & \text{otherwise.} \end{cases}$$

- **Mean**:

$$E(Y) = \frac{a+b}{2}.$$

- **Variance**:

$$\mathrm{Var}(Y) = \frac{(b-a)^2}{12}.$$

Uniform Distribution

- **Example (Inventory Delivery Timing)**: A logistics company manages deliveries for a local retail chain. Due to unpredictable but bounded road conditions, delivery trucks arrive at the warehouse uniformly at random between 8 AM and 10 AM each day. Let X be the random variable representing the arrival time of a truck, measured in hours after midnight.

- **Q1**: Identify the distribution of X. Justify your answer.

- **Q2**: Find the expected arrival time of the truck, $E(X)$.

- **Q3**: Find the variance and standard deviation of arrival times, $\mathrm{Var}(X)$ and $\mathrm{sd}(X)$.

- **Q4**: Find the probability that a truck arrives before 9 AM.

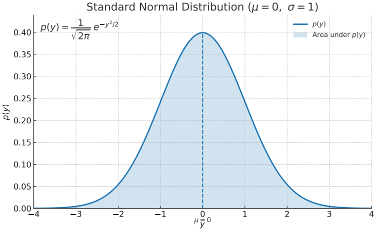
Standard Normal Random Variable

- The main reason we need to introduce continuous RVs is because we need to talk about Normal RVs.

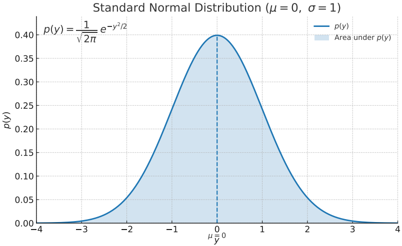
- Normal RVs are a special type of idealistic continuous random variables that are extremely important.

- **Definition**: The PDF $p(y)$ of standard normal distribution of Y is given by:

$$p(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad y \in \mathbb{R}.$$



Standard Normal Random Variable



- **Mean**:

$$E(Y) = 0.$$

- **Variance**:

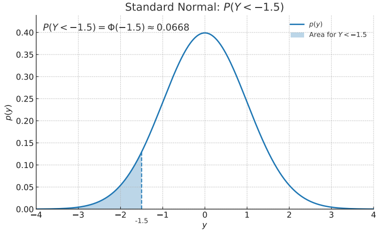
$$\mathrm{Var}(Y) = 1.$$

- This distribution is called 1) standard normal or 2) standard Gaussian or sometimes written as 3) $Y \sim N(0, 1)$.

More Practices on $Y \sim N(0,1)$

- **Example 1**: Compute $P(Y \leq -1.5)$, by the definition of PDF we can get this number by integrating the PDF:

$$P(Y < -1.5) = \int_{-\infty}^{-1.5} p(y)dy = \int_{-\infty}^{-1.5} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

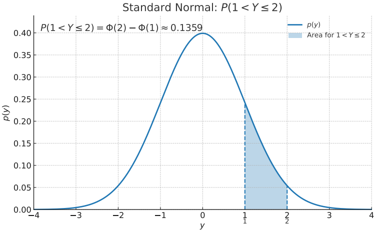


- The problem is that $e^{-y^2/2}$ is basically impossible to integrate by hand. So we need to use computers or tables, which is $P(Y < -1.5) = 0.0668$.

More Practices on $Y \sim N(0,1)$

- **Example 2**: Compute $P(1 < Y \leq 2)$, by the definition of PDF we can get this number by integrating the PDF:

$$P(1 < Y \leq 2) = \int_1^2 p(y)dy = \int_1^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

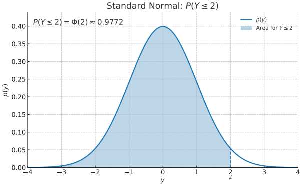


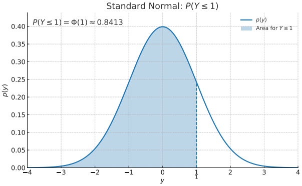
- However, table for standard Gaussian only works for areas to the left.

More Practices on $Y \sim N(0,1)$

- Here is how you can manipulate things:

$$P(1 < Y \leq 2) = P(Y \leq 2) - P(Y \leq 1).$$





- Then,

$$\begin{aligned} P(1 < Y \leq 2) &= P(Y \leq 2) - P(Y \leq 1) \\ &= 0.9772 - 0.8413 \\ &= 0.1359. \end{aligned}$$

General Normal Random Variables

- In fact, the standard normal is just one random variable in a whole class of Normal RVs:

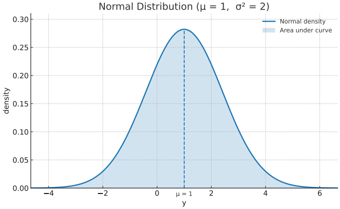
- **Definition**: For any number μ and positive numbers $\sigma^2 > 0$ the four statements are equivalent:
  (1) $X \sim N(\mu, \sigma^2)$.
  (2) X is Gaussian distribution with $E(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$.
  (3) X is Normal distribution with $E(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$.
  (4) X is a continuous random variable with PDF $p(x)$ given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

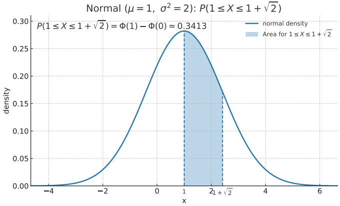
- Note that μ and σ^2 are often called parameters of the normal distribution.

General Normal Random Variables

- **Example**: if $X \sim N(1, 2)$, then $E(X) = 1$, $\mathrm{Var}(X) = \sigma^2 = 2$, and X has PDF $p(x) = \frac{1}{\sqrt{4\pi}} \exp\left(-\frac{(x-1)^2}{4}\right)$.



- **Question**: Draw the area corresponding to $P(1 \leq X \leq 1 + \sqrt{2})$.



Z-score

- **Definition**: The Z-score of a RV X is another RV Z defined as

$$Z = \frac{X - E(X)}{\mathrm{sd}(X)}.$$

- **Motivation**: We need to have score with common scale.

- **Question 1**: What is $E(Z)$?

- Note that $E(Z)$ and $\mathrm{sd}(Z)$ are considered as constants.

$$\begin{aligned} E(Z) &= E\left(\frac{X - E(X)}{\mathrm{sd}(X)}\right) = \frac{1}{\mathrm{sd}(X)} E(X - E(X)) \\ &= \frac{1}{\mathrm{sd}(X)} \{E(X) - E(E(X))\} = \frac{1}{\mathrm{sd}(X)} \{E(X) - E(X)\} = 0. \end{aligned}$$

Z-score

- **Question 2**: What is $\mathrm{sd}(Z)$?

- Note that $E(Z)$ and $\mathrm{sd}(Z)$ are considered as constants.

$$\begin{aligned}\mathrm{Var}(Z) &= \mathrm{Var}\left(\frac{X - E(X)}{\mathrm{sd}(X)}\right) = \frac{1}{(\mathrm{sd}(X))^2}\mathrm{Var}(X - E(X)) \\ &= \frac{1}{(\mathrm{sd}(X))^2}\mathrm{Var}(X) = \frac{1}{\mathrm{Var}(X)}\mathrm{Var}(X) = 1.\end{aligned}$$

Important Fact on Gaussian Random Variables

- If X is General Normal Random Variables (i.e., $X \sim N(a, b)$), then the Z-score for X is standard Gaussian, i.e.,

$$\frac{X - a}{\sqrt{b}} \sim N(0, 1).$$

General Normal Random Variables

- Using this fact allows us to compute probabilities for any Gaussian in terms of a standard Gaussian table.

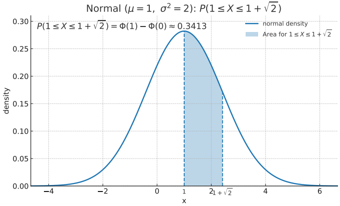
- Suppose we have $X \sim N(a, b)$ and want to find
$$P(c < X \leq d).$$

- First convert event $c < X \leq d$, so its expressed in terms of Z-score for X,
$$\frac{c-a}{\sqrt{b}} < \frac{X-a}{\sqrt{b}} \leq \frac{d-a}{\sqrt{b}},$$

- Since $\frac{X-a}{\sqrt{b}} \sim N(0, 1)$, we have the probabilities,
$$P(c < X \leq d) = P\left(\frac{c-a}{\sqrt{b}} < \frac{X-a}{\sqrt{b}} \leq \frac{d-a}{\sqrt{b}}\right)$$
$$= P(\frac{c-a}{\sqrt{b}} < Z \leq \frac{d-a}{\sqrt{b}}).$$

General Normal Random Variables

- **Question**: If $X \sim N(1,2)$, Find $P(1 \leq X \leq 1+\sqrt{2})$.



- **Answer**:

$$P(1 \leq X \leq 1+\sqrt{2}) = P\left(\frac{1-1}{\sqrt{2}} \leq \frac{X-1}{\sqrt{2}} \leq \frac{1+\sqrt{2}-1}{\sqrt{2}}\right)$$
$$= P\left(0 \leq Z \leq 1\right) = P(Z \leq 1) - P(Z < 0)$$
$$= 0.8413 - 0.5 = 0.3413.$$

Important Fact on Gaussian Random Variables

- Linear combination of Gaussian RVs is also Gaussian, then

$$W = aX + bY \sim N(c, d),$$

where $c = E(W)$ and $d = \mathrm{Var}(W)$.

- **Question**: Let $X \sim N(7.2, 2.1)$ and $Y \sim N(-1.2, 10.5)$, where $\mathrm{Cov}(X, Y) = 0$. Find $P(2X + Y \leq 10)$.

Important Fact on Gaussian Random Variables

- **Answer**: Let $W = 2X + Y$. Since W is a linear combination of Gaussian RVs, then $W \sim N(c, d)$. To find c and d,

  $$c = E(W) = E(2X + Y) = 2E(X) + E(Y) = 2 \times 7.2 - 1.2 = 13.2,$$

  and

  $$\begin{aligned} d = \mathrm{Var}(W) &= \mathrm{Var}(2X + Y) = 4\mathrm{Var}(X) + \mathrm{Var}(Y) + 4\mathrm{Cov}(X, Y) \\ &= 4 \times 2.1 + 10.5 + 0 = 18.9. \end{aligned}$$

  Then, $W \sim N(13.2, 18.9)$. To find the probability,

  $$\begin{aligned} P(2X + Y \leq 10) &= P(W \leq 10) \\ &= P\left(\frac{W - 13.2}{\sqrt{18.9}} \leq \frac{10 - 13.2}{\sqrt{18.9}}\right) \\ &= P(Z \leq -0.73) = 0.2327. \end{aligned}$$