Week 6: Normal Distribution and Expectations of Continuous RVs

MATH697

Sahir Bhatnagar

October 10, 2017

McGill University

4. The Normal Distribution

Bell-Shaped Curve

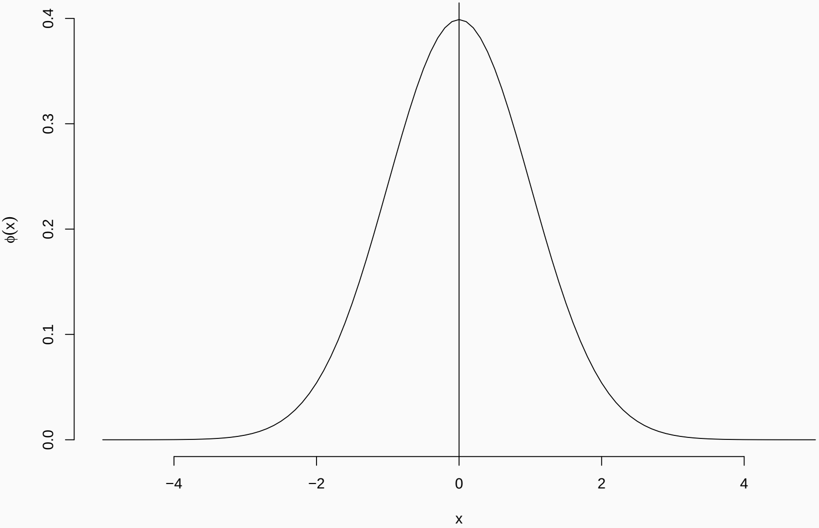
We now define a function $\phi : \mathbb{R} \to \mathbb{R}$ by

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad -\infty < x < \infty \tag{1}$$

- This function ϕ is the famous bell-shaped curve because its graph is in the shape of a bell.

Bell-Shaped Curve

```r
curve(dnorm(x), from = -5, to = 5, ylab = expression(phi(x)), bty = "n")
abline(v=0)
```



Example 1 (Bell-Shaped Curve)

Verify that Equation (1) is really a density function.

Proof: on board. note, this is an example of an integration that either you know how to do, or else you can spend a very long time going nowhere.

A probability model that reflects observed (**empirical**) behaviour of data samples; this distribution is often observed in practice:

Definition 2 (The Standard Normal Distribution)

Let $X \sim N(0,1)$. Then X has density function

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad -\infty < x < \infty \tag{2}$$

The Standard Normal Distribution

$X \sim N(0,1)$. This means that for $-\infty < a \leq b < \infty$,

$$P(a \leq X \leq b) = \int_a^b \phi(x)dx = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx$$

Let $\mu \in \mathbb{R}$ and $\sigma > 0$.

Definition 3 (The Normal(μ, σ^2) Distribution)

$$\begin{aligned} f_X(x) &= \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad x \in \mathbb{R}. \end{aligned} \tag{3}$$

Let X be a random variable having density function given by (3). The RV X is said to have the $N(\mu, \sigma^2)$ distribution. We write this as $X \sim N(\mu, \sigma^2)$

Normal Distribution is a density function

Example 4 (Normal Distribution)

Verify that Equation (3) is really a density function.

Proof: on board.

The Normal Distribution

The pdf is symmetric about μ, and hence μ is controls the *location* of the distribution and σ^2 controls the *spread* or *scale* of the distribution.

1. The Normal density function is justified by the **Central Limit Theorem**.

The pdf is symmetric about μ, and hence μ is controls the *location* of the distribution and σ^2 controls the *spread* or *scale* of the distribution.

1. The Normal density function is justified by the **Central Limit Theorem**.
2. Special case: $\mu = 0, \sigma^2 = 1 \rightarrow$ the **standard** or **unit** normal distribution. In this case, the density function is denoted $\phi(x)$, and the cdf is denoted $\Phi(x)$:

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt = \int_{-\infty}^{x} \left(\frac{1}{2\pi}\right)^{1/2} \exp\left\{-\frac{1}{2}t^2\right\}\, dt.$$

The pdf is symmetric about μ, and hence μ is controls the *location* of the distribution and σ^2 controls the *spread* or *scale* of the distribution.

1. The Normal density function is justified by the **Central Limit Theorem**.
2. Special case: $\mu = 0, \sigma^2 = 1 \rightarrow$ the **standard** or **unit** normal distribution. In this case, the density function is denoted $\phi(x)$, and the cdf is denoted $\Phi(x)$:

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt = \int_{-\infty}^{x} \left(\frac{1}{2\pi}\right)^{1/2} \exp\left\{-\frac{1}{2}t^2\right\}\, dt.$$

The Normal Distribution

The pdf is symmetric about μ, and hence μ is controls the *location* of the distribution and σ^2 controls the *spread* or *scale* of the distribution.

1. The Normal density function is justified by the **Central Limit Theorem**.
2. Special case: $\mu = 0, \sigma^2 = 1 \rightarrow$ the **standard** or **unit** normal distribution. In this case, the density function is denoted $\phi(x)$, and the cdf is denoted $\Phi(x)$:

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt = \int_{-\infty}^{x} \left(\frac{1}{2\pi}\right)^{1/2} \exp\left\{-\frac{1}{2}t^2\right\}\, dt.$$

This integral can only be calculated numerically.

The Normal Distribution

3. If $X \sim N(\mu, \sigma^2)$, and $Z = (X - \mu)/\sigma$, then $Z \sim N(0, 1)$, *Proof*: on board

3. If $X \sim N(\mu, \sigma^2)$, and $Z = (X - \mu)/\sigma$, then $Z \sim N(0, 1)$, *Proof*: on board
4. If $X \sim N(0, 1)$, and $Y = \sigma X + \mu$, then $Y \sim N(\mu, \sigma^2)$, *Proof*: on board

The Normal Distribution

3. If $X \sim N(\mu, \sigma^2)$, and $Z = (X - \mu)/\sigma$, then $Z \sim N(0, 1)$, *Proof*: on board
4. If $X \sim N(0, 1)$, and $Y = \sigma X + \mu$, then $Y \sim N(\mu, \sigma^2)$, *Proof*: on board
5. If $X \sim N(0, 1)$, and $Y = X^2$, then $Y \sim Gamma(1/2, 1/2) = \chi^2_1$.

The Normal Distribution

3. If $X \sim N(\mu, \sigma^2)$, and $Z = (X - \mu)/\sigma$, then $Z \sim N(0, 1)$, *Proof*: on board
4. If $X \sim N(0, 1)$, and $Y = \sigma X + \mu$, then $Y \sim N(\mu, \sigma^2)$, *Proof*: on board
5. If $X \sim N(0, 1)$, and $Y = X^2$, then $Y \sim Gamma(1/2, 1/2) = \chi^2_1$.

3. If $X \sim N(\mu, \sigma^2)$, and $Z = (X - \mu)/\sigma$, then $Z \sim N(0, 1)$, *Proof*: on board
4. If $X \sim N(0, 1)$, and $Y = \sigma X + \mu$, then $Y \sim N(\mu, \sigma^2)$, *Proof*: on board
5. If $X \sim N(0, 1)$, and $Y = X^2$, then $Y \sim Gamma(1/2, 1/2) = \chi^2_1$.

6. If $X \sim N(0, 1)$ and $Y \sim \chi^2_\alpha$ are independent random variables, then random variable T, defined by

$$T = \frac{X}{\sqrt{Y/\alpha}}$$

has a **Student-t distribution** with α **degrees of freedom**. The Student-t distribution plays an important role in certain statistical testing procedures.

PDF and CDF of N(0,1) (Standard Normal)

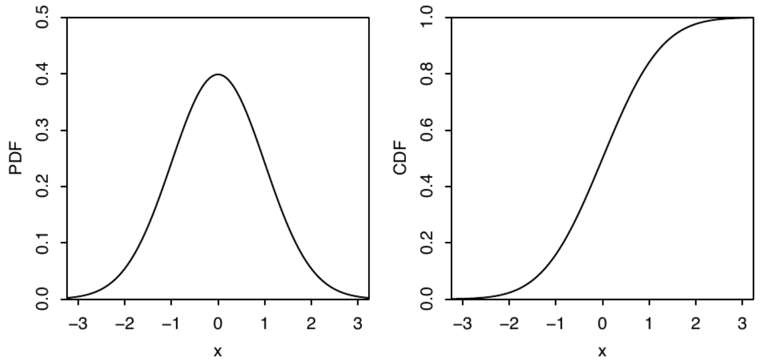


FIGURE 5.9
Standard Normal PDF φ (left) and CDF Φ (right).

Figure 1

Important Symmetry Properties of Standard Normal

1. *Symmetry of PDF*: ϕ statisfies $\phi(z) = \phi(-z)$, i.e., ϕ is an even function

Important Symmetry Properties of Standard Normal

1. *Symmetry of PDF*: ϕ statisfies $\phi(z) = \phi(-z)$, i.e., ϕ is an even function
2. *Symmetry of Tail Areas*: The area under the PDF curve to the left of -2 which is $P(Z \leq -2) = \Phi(-2)$ by definition, equals the area to the right of 2, which is $P(Z \geq 2) = 1 - \Phi(2)$. In general,

$$\Phi(z) = 1 - \Phi(-z)$$

Important Symmetry Properties of Standard Normal

1. *Symmetry of PDF*: ϕ statisfies $\phi(z) = \phi(-z)$, i.e., ϕ is an even function
2. *Symmetry of Tail Areas*: The area under the PDF curve to the left of -2 which is $P(Z \leq -2) = \Phi(-2)$ by definition, equals the area to the right of 2, which is $P(Z \geq 2) = 1 - \Phi(2)$. In general,

$$\Phi(z) = 1 - \Phi(-z)$$

3. *Symmetry of Z and $-Z$*: if $Z \sim N(0, 1)$, then $-Z \sim N(0, 1)$ as well. To see this not that the CDF of $-Z$ is

$$P(-Z \leq z) = P(Z \geq -z) = 1 - \Phi(-z) = \Phi(z)$$

Exercise in R: Normal Distribution

Exercise 5 (Normal Distribution: Effect of σ and μ)

1. *Plot the Normal densities for* $\mu = 0$ *and* $\sigma^2 = 1, 4, 9, 16$ *on the same plot in different colors. Add a legend.*
2. *Plot the Normal densities for* $\mu = 0, 1, 2, 3, 4$ *and* $\sigma^2 = 1$ *on the same plot in different colors. Add a legend.*

Reading Probabilities of the Table

The $N(0,1)$ table that will be used in this course provides the following area:

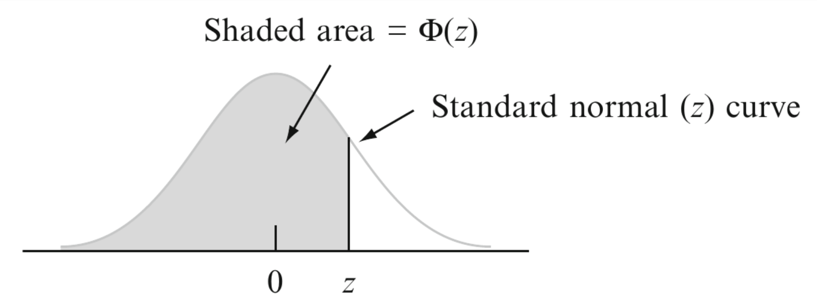


Figure 2

Reading Probabilities of the Table

Exercise 6 (Normal Distribution Table)

Suppose $Z \sim N(0,1)$. Find the following using the table.

1. $P(0 \leq Z \leq 1.4)$
2. $P(0 \leq Z \leq 1.42)$
3. $P(Z > 1.42)$
4. $P(Z < -1.42)$
5. $P(-1.5 < Z < 1.42)$
6. $P(1.25 < Z < 1.42)$
7. *Confirm your results in* R

Percentiles of the Standard Normal Distribution

For any p between 0 and 1, the standard normal table can be used to obtain the $(100p)$th percentile of the standard normal distribution.

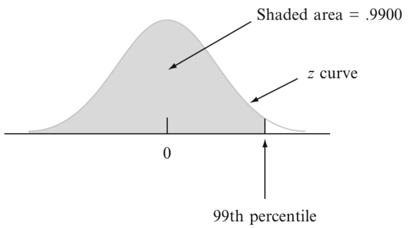
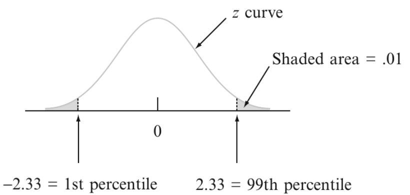
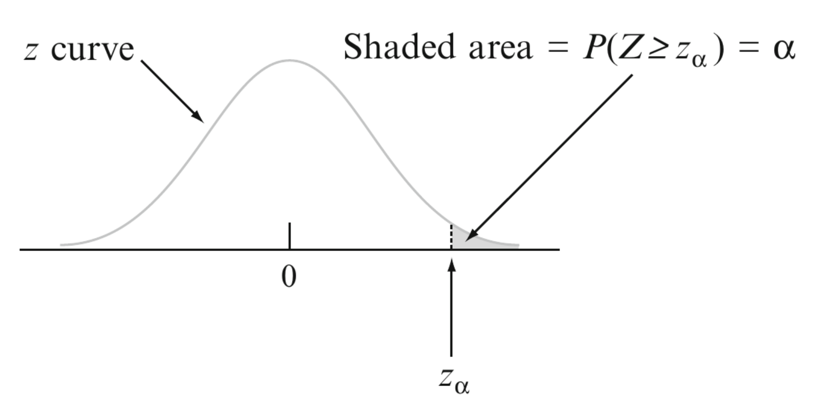


Figure 4.17 **Finding the 99th percentile**



In statistical inference, we will need the values on the measurement axis that capture certain small tail areas under the standard normal curve. z_α will denote the value on the measurement axis for which α of the area under the z curve lies to the right of z_α.



Example 7 (95th Percentile)

The $100(1-0.05)$th $=$ 95th percentile of the standard normal distribution is $z_{.05}$, so $z_{.05} = 1.645$. The area under the standard normal curve to the left of $-z_{.05}$ is also .05.

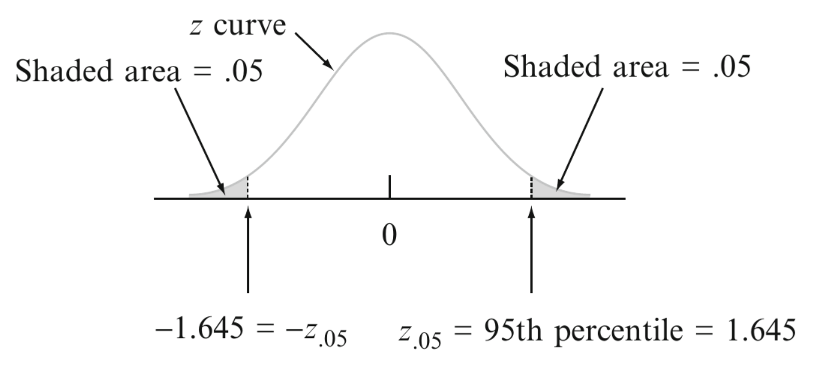


Figure 5

Exercise 8 (Normal Distribution Table)

Suppose $Z \sim N(0, 1)$. Find z so that

1. $P(Z > z) = 0.05$
2. $P(Z > z) = 0.025$
3. *Confirm your results in* R

Exercise 9 (Normal Distribution Table)

Suppose $X \sim N(-2, 9)$. *Find*

1. $P(-6.5 < X < 2.26)$
2. *Confirm your result in* R

Expectation and Variance for Continous RVs

- We introduced the distribution of a RV, which gives us full information about the probability that the RV will fall into any particular set.

Motivation for Expectation and Variance

- We introduced the distribution of a RV, which gives us full information about the probability that the RV will fall into any particular set.
- For example, we can say how likely it is that the RV will exceed 1000, that it will equal 5, and that it will be in the interval [0, 7]

Motivation for Expectation and Variance

- We introduced the distribution of a RV, which gives us full information about the probability that the RV will fall into any particular set.
- For example, we can say how likely it is that the RV will exceed 1000, that it will equal 5, and that it will be in the interval $[0, 7]$
- It can be difficult to manage so many probabilities though, so often we want just one number summarizing the "average" value of the RV $\longrightarrow$ Expected Value

Motivation for Expectation and Variance

- We introduced the distribution of a RV, which gives us full information about the probability that the RV will fall into any particular set.
- For example, we can say how likely it is that the RV will exceed 1000, that it will equal 5, and that it will be in the interval $[0, 7]$
- It can be difficult to manage so many probabilities though, so often we want just one number summarizing the "average" value of the RV $\longrightarrow$ Expected Value
- In addition, much of statistics is about understanding variability in the world, so it is often important to know how spread out the distribution is $\longrightarrow$ Variance

Expectation and Variance

Recall for a **discrete** random variable X taking values in set $\mathbb{X}$ with mass function f_X, the **expectation** of X is defined by

$$E_{f_X}[\, X\,] = \sum_{x \in \mathbb{X}} x f_X(x) \equiv \sum_{x=-\infty}^{\infty} x f_X(x)$$

Recall for a **discrete** random variable X taking values in set $\mathbb{X}$ with mass function f_X, the **expectation** of X is defined by

$$E_{f_X}[\,X\,] = \sum_{x \in \mathbb{X}} x f_X(x) \equiv \sum_{x=-\infty}^{\infty} x f_X(x)$$

For a **continuous** random variable X with pdf f_X, the expectation of X is defined by

$$E_{f_X}[X] = \int_{\mathbb{X}} x f_X(x)\, dx \equiv \int_{-\infty}^{\infty} x f_X(x)\, dx$$

Expectation and Variance

Recall for a **discrete** random variable X taking values in set $\mathbb{X}$ with mass function f_X, the **expectation** of X is defined by

$$E_{f_X}[\,X\,] = \sum_{x \in \mathbb{X}} x f_X(x) \equiv \sum_{x=-\infty}^{\infty} x f_X(x)$$

For a **continuous** random variable X with pdf f_X, the expectation of X is defined by

$$E_{f_X}[X] = \int_{\mathbb{X}} x f_X(x)\,dx \equiv \int_{-\infty}^{\infty} x f_X(x)\,dx$$

The **variance** of X is defined by

$$Var_{f_X}[X] = E_{f_X}[(X - E_{f_X}[X])^2] = E_{f_X}\left[X^2\right] - \left\{E_{f_X}[X]\right\}^2.$$

Expectation and Variance Interpretation

The expectation and variance of a probability distribution can be used to aid description, or to characterize the distribution; the EXPECTATION is a measure of **location** (that is, the "centre of mass" of the probability distribution. The VARIANCE is a measure of **scale** or **spread** of the distribution (how widely the probability is distributed) .

Center of mass of two pebbles, depicting that $E(X) = p$ for $X \sim Bernoulli(p)$. Here q and p denote the masses of the two pebbles.

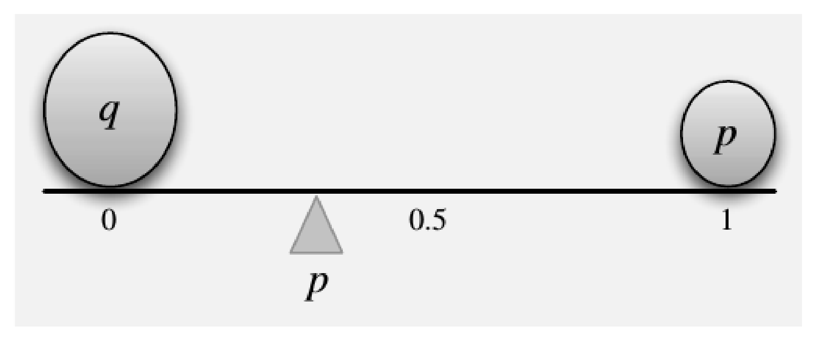
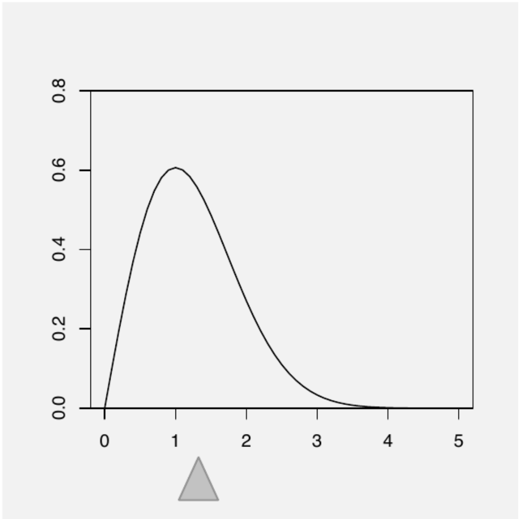


Figure 6

The expected value of a continuous RV is the balancing point of the PDF.



Expectation Example Continuous RV

Example 10 (A continuous RV)

Suppose that X is a continuous random variable taking values on $\mathbb{X} = \mathbb{R}^+$ with pdf

$$f_X(x) = \frac{2}{(1+x)^3} \qquad x > 0.$$

Then, integrating by parts.

$$\begin{aligned} E_{f_X}[X] &= \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^{\infty} \frac{2x}{(1+x)^3}\,dx = \left[-\frac{x}{(1+x)^2}\right]_0^{\infty} + \int_0^{\infty} \frac{1}{(1+x)^2}\,dx \\ &= 0 - \left[-\frac{1}{1+x}\right]_0^{\infty} = 1 \end{aligned}$$

Expectations of Sums of RVs

Suppose that X_1 and X_2 are independent random variables, and a_1 and a_2 are constants. Then if $Y = a_1X_1 + a_2X_2$, it can be shown that

$$\begin{aligned} E_{f_Y}[Y] &= a_1E_{f_{X_1}}[X_1] + a_2E_{f_{X_2}}[X_2] \\ Var_{f_Y}[Y] &= a_1^2Var_{f_{X_1}}[X_1] + a_2^2Var_{f_{X_2}}[X_2] \end{aligned}$$

so that, in particular (when $a_1 = a_2 = 1$) we have

$$\begin{aligned} E_{f_Y}[Y] &= E_{f_{X_1}}[X_1] + E_{f_{X_2}}[X_2] \\ Var_{f_Y}[Y] &= Var_{f_{X_1}}[X_1] + Var_{f_{X_2}}[X_2] \end{aligned}$$

so we have a simple additive property for expectations and variances.

Sums of RVs

Sums of random variables crop up naturally in many statistical calculations. Often we are interested in a random variable Y that is defined as the sum of some other **independent and identically distributed** (i.i.d) random variables, $X_1, \ldots, X_n$. If

$$Y = \sum_{i=1}^{n} X_i \qquad \textit{with} \qquad E_{f_{X_i}}[X_i] = \mu \quad \textit{and} \quad Var_{f_{X_i}}[X_i] = \sigma^2$$

we have

$$E_{f_Y}[Y] = \sum_{i=1}^{n} E_{f_{X_i}}[X_i] = \sum_{i=1}^{n} \mu = n\mu$$

and

$$Var_{f_Y}[Y] = \sum_{i=1}^{n} Var_{f_{X_i}}[X_i] = \sum_{i=1}^{n} \sigma^2 = n\sigma^2$$

and also, if

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{is the } \textbf{sample mean} \text{ random variable}$$

then, using the properties listed above

$$E_{f_{\bar{X}}}\left[\bar{X}\right] = \frac{1}{n} E_{f_Y}[Y] = \frac{1}{n} n\mu = \mu$$

and

$$Var_{f_Y}[Y] = \frac{1}{n^2} Var_{f_Y}[Y] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Expectations of a function of a RV

Suppose that X is a random variable, and $g(.)$ is some function. Then we can define the expectation of $g(X)$ (that is, the expectation of a function of a random variable) by

$$E_{f_X}[\, g(X)\,] = \begin{cases} \displaystyle\sum_{x=-\infty}^{\infty} g(x)f_X(x) & \text{DISCRETE CASE} \\[2ex] \displaystyle\int_{-\infty}^{\infty} g(x)f_X(x)\; dx & \text{CONTINUOUS CASE} \end{cases}$$

Expectations of a function of a RV

For example, if X is a continuous random variable, and $g(x) = \exp\{-x\}$ then

$$E_{f_X}[\, g(X)\,] = E_{f_X}[\, \exp\{-X\}\,] = \int_{-\infty}^{\infty} \exp\{-x\}\, f_X(x)\, dx$$

Note that $Y = g(X)$ is also a random variable whose probability distribution we can calculate from the probability distribution of X.

Mean of an Exponential RV

Recall that if $X \sim \textit{Exponential}(\lambda)$, then $f_X(x) = \lambda e^{-\lambda x}$, for $x \geq 0$. The λ parameter is known as the rate paramter. Show that $E(X) = 1/\lambda$.

Proof: on board.