MTH 3210 Lab 2

Due Thu., Mar. 21

1 Part A: Discrete Probability Distributions

1.1 The Binomial Distribution

We use the function `dbinom()` to compute a **probability**

$$P(X = x) \; = \; \binom{n}{x} p^x (1-p)^{n-x}$$

from a **binomial(n, p)** distribution.

The `dbinom()` function takes arguments:

`x`	The value x in $P(X = x)$, which can be a vector.
`size`	the number of trials n.
`prob`	the success probability p.

It returns the value of $P(X = x)$. Type `?dbinom` for more information.

We use the function `pbinom()` to compute the **cumulative probability**

$$P(X \leq x) \; = \; \sum_{y=0}^{x} \binom{n}{y} p^y (1-p)^{n-y} \,.$$

The `pbinom()` function takes arguments:

`q`	The value x in $P(X \leq x)$, which can be a vector.
`size`	the number of trials n
`prob`	the success probability p

It returns the value of $P(X \leq x)$.

1. Use `dbinom()` to compute $P(X = 7)$, where $\boldsymbol{X \sim}$ **binom(10, 0.2)**.
2. Now use `dbinom()` to compute $P(3 \leq X \leq 7)$.
3. Use `pbinom()` to compute $P(X \leq 7)$.
4. Now use `pbinom()` to compute $P(3 \leq X \leq 7)$.

1.2 The Geometric Distribution

We use the function `dgeom()` to compute a probability

$$P(X = x) \;=\; (1-p)^{x-1}p$$

from a **geometric(p)** distribution.

The `dgeom()` function takes arguments:

`x`	The value x in $P(X = x)$, which can be a vector.
`prob`	the success probability p.

It returns the value of $P(X = x)$.

We use the function `pgeom()` to compute the **cumulative probability**

$$P(X \le x) \;=\; \sum_{y=0}^{x}(1-p)^{y-1}p\,.$$

The `pgeom()` function takes arguments:

`q`	the value x in $P(X \le x)$, which can be a vector.
`prob`	the success probability p

It returns the value of $P(X \le x)$.

Note that in R, a geometric random variable Y is defined to be the number of *failures prior* to obtaining the first success (as opposed to the number of *failures plus the final success*), i.e. $\boldsymbol{Y = X - 1}$.

1. Use `dgeom()` to compute $P(X = 5)$, where $\boldsymbol{X \sim}$ **geometric(0.2)**.
2. Use `pgeom()` to compute $P(X \le 5)$.
3. Now use `pgeom()` to compute $P(3 \le X \le 5)$.

2 Part B: Continuous Probability Distributions

2.1 The Uniform Distribution

We use the function `punif()` to compute the **cumulative probability**

$$P(X \le x) \;=\; \int_A^x \frac{1}{A+B}\,dy \qquad\qquad (\text{for } A \le x \le B)$$

from a **uniform(A, B)** distribution.

The `punif()` function takes arguments:

`q`	the value x in $P(X \le x)$, which can be a vector.
`min`	the lower endpoint A of the interval over which X ranges
`max`	the upper endpoint B of the interval over which X ranges

It returns the value of $P(X \leq x)$.

We use the function `runif()` to generate a **random sample** from a **uniform($\boldsymbol{A, B}$)**.

The `runif()` function takes arguments:

`n`	the sample size.
`min`	the lower endpoint A of the interval over which X ranges
`max`	the upper endpoint B of the interval over which X ranges

It returns a vector containing the **random sample**.

1. Use `punif()` to compute $P(X \leq 5.5)$, where $\boldsymbol{X \sim}$ **uniform($\boldsymbol{0, 10}$)**.

2. Now use `punif()` to compute $P(4.5 \leq X \leq 5.5)$.

3. Use `runif()` to generate a random sample of size $n = 10,000$ from the uniform$(0, 10)$ distribution, and use `hist()` to make a histogram of the simulated sample values. How does the histogram compare to the **uniform($\boldsymbol{0, 10}$) pdf**?

4. Recall that if $X \sim$ uniform(A, B), then

   $$E(X) \;=\; (A + B)/2$$

   and

   $$SD(X) \;=\; \sqrt{(B - A)^2/12}$$

   Compute $\boldsymbol{E(X)}$ and $\boldsymbol{SD(X)}$ for the **uniform($\boldsymbol{0, 10}$)** distribution.

   Then use `mean()` and `sd()` to compute the **sample mean** and **sample standard deviation** of your $n = 10,000$ simulated **uniform($\boldsymbol{0, 10}$)** values. How do their values compare to the theoretical mean $E(X)$ and standard deviation $SD(X)$?