# Lab 4: Probability Distributions

*STAT 111*

*6/11/2015*

**Complete the following problems on a separate piece of paper. Read these instructions carefully.**

- Hand in your answers at the end of lab – make sure to include you name and the lab number. Instructions are included for making graphs and finding the necessary statistics for each problem.
- Each question will be worth two points, and you can receive partial credit for incorrect answers if your process was correct.
- Write your final answer **as a sentence** and include all steps you used to get there, otherwise you will receive partial credit.
- When performing calculations, keep intermediate steps rounded to **four decimal places**, and round your final answer to **three decimal places**.

## Formulas:

| Distribution | $E(X)$ | $Var(X)$ | $SD(X)$ |
|---|---|---|---|
| $X \sim Bernoulli(p)$ | $p$ | $p(1-p)$ | $\sqrt{p(1-p)}$ |
| $X \sim Binomal(n, p)$ | $np$ | $np(1-p)$ | $\sqrt{np(1-p)}$ |
| $X \sim Poisson(\lambda)$ | $\lambda$ | $\lambda$ | $\sqrt{\lambda}$ |

## In StatCrunch:

For unnamed distributions:

- Enter the table with the possible outcomes as one column and their corresponding probabilities as another column.
- `Stat` → `Calculators` → `Custom`
- Select $X$ column as `Values in`, Select $P(x)$ column as `Weights in`
- Hit `Compute!` to find $E(X)$ and $SD(X)$

For named distributions:

- `Stat` → `Calculators` → Choose the distribution
- Enter the relevant parameters ($n$, $p$, $\lambda$, $mu$, $\sigma$, etc.)
- Enter range of interest, e.g. $P(X \leq 20)$
- Hit `Compute!`

## Part A: The Lottery

Say you play a scratch-off lottery ticket which costs \$10 to play. You get three numbers, and the prize tiles are randomly numbered. Your prize is determined by how many of your numbers show up in the prize tiles, with payouts and probabilities determined as follows:

| Outcome | Prize ($X$) | Probability $P(X = x)$ |
|---|---|---|
| Match all Three | \$1,000 | 0.001 |
| Match Two | \$100 | 0.01 |
| Match One | Free Ticket (\$10) | 0.1 |
| Match None | \$0 | 0.889 |

1. What is your expected prize?

$E(X) = \$3$, so we expect to win \$3 on any given game, or average \$3 in winnings if we play many times.

2. What is the standard deviation of the prizes?

From StatCrunch, $SD(X) \approx \$33.18$. The standard deviation of prize winnings is \$33.18.

3. How much should you expect to win (or lose) on a single ticket? Hint: Winnings = Prize - Cost

The game costs \$10 and we expect to win \$3 per game in the long term. Our expected winnings are \$3 - \$10 = -\$7, so we should expect to lose \$7 per game in the long term.

## Part B: Multiple Choice

Assume you need to take a multiple choice statistics quiz with ten questions. You haven't studied or attended lectures, so you guess the answer to each question completely at random. Each question has four choices.

1. For a single question, what is the probability of guessing the right answer?

There are four choices for every question, and only one of them is correct. Since all choices are equally likely, $P(C) = \frac{1}{4} = 0.25$.

2. What is the probability distribution for the number of correct answers on the **entire** quiz? Make sure to include the name and any necessary parameters.

There are 10 questions, and each is a Bernoulli trial with $p = 0.25$. The total number of correct answer is the sum of these 10 trials, so $X \sim Binomial(n = 10, p = 0.25)$.

3. What are the mean and standard deviation for the distribution of the number of correct answers?

$$E(X) = np = 10(0.25) = 2.5$$
$$SD(X) = \sqrt{np(1-p)} = \sqrt{10(0.25)(0.75)} = \sqrt{1.875} \approx 1.369$$

We should expect to get 2.5 questions correct, with a standard deviation of 1.369 questions.

4. Say you need answer at least 60% of questions correctly to pass the quiz. What is the probability of this happening?

Earning at least a 60% means getting at least 6 questions correct. From StatCrunch, $P(X \geq 6) = 0.0197$, so there is only a 1.97% chance of this happening.

## Part C: Number of Forest Fires

Say that Alaska averages 3 lost hikers per month in any given summer.

1. What is the probability distribution of lost hikers in a month? Make sure to include the name and any necessary parameters.

We have an average number of occurances and a unit of time, so $X \sim Poisson(\lambda = 3)$.

2. What is the distribution of lost hikers for an entire summer (assume that a summer is 3 months). Make sure to include the name and any necessary parameters.

Since we expect 3 hikers to get lost per month and there are 3 months in a summer, we should expect to have 9 lost hikers in a given summer, so $Y \sim Poisson(\lambda = 9)$.

3. What is the probability that fewer than 5 hikers get lost in an entire summer?

From StatCrunch, $P(Y < 5) = 0.055$. There is only a 5.5% chance that we lose fewer than 5 hikers.