MIT Introduction to Statistics 18.05 Class 7 Slides - Solutions

John Hancock

April 17, 2017

Contents

1 References and License

We are answering questions in the material from MIT OpenCourseWare course 18.05, Introduction to Probability and Statistics.
In this document we are answering questions Orloff and Bloom ask in [6].
Please see the references section for detailed citation information.
The material for the course is licensed under the terms at `http://ocw.mit.edu/terms`.
We use documentation in [1], [7] to write LaTeXsource code for this document.

2 Estimate Error

The first question Orloff and Bloom ask in [6] is about an accountant that rounds his cacluations (entries) to the nearest dollar. We assume the accountant has made 300 calculations. Orloff and Bloom want us to estimate the probability that the total error is greater than five dollars.
We use the central limit theorem [3] and techniques for estimating probability that Orloff and Bloom show in [3] in order to find this estimate.
In order to apply the central limit theorem, we first define a random variable X_i. X_i is the error the accountant makes on her i^th calculation. Orloff and Bloom tell us that X_i is uniformly distributed on $[-0.5, 0.5]$.
We also need the mean μ, and standard deviation σ of X_i in order to make our estimate.

In [5] Orloff and Bloom state that a uniformly distributed random variable on $[a, b]$ has the distribution function $f(x) = \frac{1}{a-b}$.
In [4] Orloff and Bloom define the mean $E(X)$ of a continuous random variable X with pdf $f(x)$ to be:

$$E(X) = \int_a^b xf(x)\,dx \tag{1}$$

For this problem, $f(x) = \frac{1}{-0.5-0.5} = -1$. Therefore we apply Orloff and Bloom's definition of the mean value of a continuous random variable to find that the mean value of X_i is

$$E(X_i) = \int_{-0.5}^{0.5} -x\,dx. \tag{2}$$

We use the power rule for integrals from [2] to find the antiderivative of the function above that we must integrate in order to find the mean value of X_i.
The antiderivative of $g(x) = -x$ is $\frac{-x^2}{2}$.
We replace the integral on the right hand side of the equation above with its antiderivative:

$$E(X_i) = \left.\frac{-x^2}{2}\right|_{-0.5}^{0.5}. \tag{3}$$

And we evaluate the antiderivative over the interval $[-0.5, 0.5]$:

$$E(X_i) = \frac{-\left(-0.5^2\right)}{2} - \frac{-\left(0.5^2\right)}{2} \tag{4}$$

Now we do some arithmetic to simplify the right hand side of the equation above:

$$E(X_i) = \frac{-1}{8} - \frac{-1}{8} = 0. \tag{5}$$

In order to find the standard deviation of X_i, we use a property of variance from [4], for a continuous random variable X:

$$\mathrm{Var}(X) = E\left(X^2\right) - E(X)^2. \tag{6}$$

We apply the same reasoning to find $E\left(X_i^2\right)$ that we use to find $E(X_i)$:

$$E\left(X_i^2\right) = \int_{-0.5}^{0.5} -\left(x^2\right)\,dx. \tag{7}$$

This implies:

$$E\left(X_i^2\right) = \left.\frac{-x^3}{3}\right|_{-0.5}^{0.5}. \tag{8}$$

Which implies

$$E\left(X_i^2\right) = \frac{-\left(-0.5^3\right)}{3} - \frac{-\left(0.5^3\right)}{3} \tag{9}$$

The right hand side of the equation above simplifies to:

$$E\left(X_i^2\right) = \frac{-\left(-1\right)}{24} - \frac{-1}{24} \tag{10}$$

Therefore the variance of X_i is $\frac{1}{12}$.
Now we define a random variable S to be the sum of 300 values of the X_i. Therefore S is the total erorr that the accountant makes after 300 calculations.
The mean value $E\left(S\right)$
We standardize S, and apply the central limit theorem like Orloff and Bloom do in [3] to get the approximation:

$$P\left(\frac{S-0}{300\sqrt{\frac{1}{12}}} > \frac{5-0}{300\sqrt{\frac{1}{12}}}\right) \approx P\left(Z > \frac{5}{300\sqrt{\frac{1}{12}}}\right) \tag{11}$$

We use python to approximate $\frac{5}{300\sqrt{\frac{1}{12}}} \approx 0.0577$.
Now we use the pnorm function of R to find the probability that $Z > 0.0577$. The value of R's pnorm function for the input value 0.0577 is approximately 0.523.
Therefore the probability that the total error the account makes after 300 calculations is 0.523.

References

[1] Latex.org user alainremillard. *Logical Not Symbol.* Available at `http://www.math.csusb.edu/notes/logic/lognot/node2.html` (2012/3/16).

[2] Michael Dougherty. *Chapter 6 Basic Integration.*

[3] Jeremy Orloff and Jonathan Bloom. *Expectation, Variance and Standard Deviation for Continuous Random Variables Class 6, 18.05, Spring 2014 Jeremy Orloff and Jonathan Bloom.* Available at `https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading6b.pdf` (Spring 2014).

[4] Jeremy Orloff and Jonathan Bloom. *Expectation, Variance and Standard Deviation for Continuous Random Variables Class 6, 18.05, Spring 2014 Jeremy Orloff and Jonathan Bloom*. Available at https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading6a.pdf (Spring 2014).

[5] Jeremy Orloff and Jonathan Bloom. *Gallery of Continuous Random Variables Class 5, 18.05, Spring 2014 Jeremy Orloff and Jonathan Bloom*. Available at https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading5c.pdf (Spring 2014).

[6] Jeremy Orloff and Jonathan Bloom. *Joint Distributions, Independence Covariance and Correlation 18.05 Spring 2014 Jeremy Orloff and Jonathan Bloom*. Available at https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/class-slides/MIT18_05S14_class7slides.pdf (Spring 2014).

[7] ShareLatex.com. *Theorems and proofs*. Available at https://www.sharelatex.com/learn/Theorems_and_proofs (2017).