PLSC 502: “Statistical Methods for Political Research”

Exercise Three
October 7, 2024

Introduction

The purpose of this exercise is twofold: to familiarize you with the various distributions we discussed in class, and to develop your skills at generating, manipulating, and describing (graphically and in words) random variables using statistical software. There are no “real” data for this exercise; it entirely comprises simulations. In writing up your homework, be sure to include all the code necessary to replicate your work *exactly*.

Exercise

1. **Binomial**

   (a) Generate 100 random draws from a binomial(10,0.8) distribution. Briefly describe (graphically and in words) the resulting “data,” and compare them to the theoretical distribution of values from that distribution.

   (b) Generate another 100 random draws, this time from a binomial(10,0.2) density. Describe these “data,” and compare them to those in part (a).

   (c) Generate 50,000 draws from a binomial(10,0.8) density, and compare the results both to those above and to the theoretical distribution of values. What – if anything – is different, and why?

2. **Poisson**

   (a) Generate 1,000 random draws from a Poisson distribution with $\lambda = 0.5$. Compare the proportions observed to the theoretical values.

   (b) Do the same 1,000 draws from Poisson densities with $\lambda = 1.0$, $\lambda = 5.0$, and $\lambda = 10.0$. Discuss how the observed distributions of values change as λ increases.

3. **Normal / Chi-Square / *t***

   (a) Draw 5,000 random $N(-2, 2)$ values, 5,000 $N(5, 2)$ values, and 5,000 $N(-2, 9)$ values. Plot and discuss the differences between the three sets of values.

   (b) Using random draws (however many you choose) of a standard normal variate Z, illustrate graphically that $Z^2 \sim \chi^2_1$.

   (c) Similarly, illustrate that $W_3 = Z_1^2 + Z_2^2 + Z_3^2 \sim \chi^2_3$.

   (d) Build on (b) and (c) to show that $\frac{Z}{W_3/3} \sim t_3$.

4. A *Gumbel* distribution is a two-parameter ($\alpha \in \mathbb{R}, \beta > 0$) distribution with

$$f(x) = \frac{1}{\beta} \exp\left[\frac{-(x-\alpha)}{\beta}\right] \times \exp\left\{-\exp\left[\frac{-(x-\alpha)}{\beta}\right]\right\}$$

and

$$F(x) = 1 - \exp\left\{-\exp\left[\frac{-(x-\alpha)}{\beta}\right]\right\}.$$

It's also sometimes called the "Type I Extreme Value" distribution. It is related to the standard uniform distribution by $X = \alpha - \beta \ln[-\ln(U_{0,1})]$.[1]

(a) Generate 5,000 draws from a Gumbel(1,2) distribution *via transformation of a standard uniform variate*.

(b) Plot those values, and discuss/describe the resulting empirical distribution as compared to the theoretical density.

Use plots, words, or combinations thereof to complete this exercise. Submit your answers **in PDF format**. In addition to your answers, please include a copy of all computer code used to conduct your simulations, generate your figures, etc. This can be in any form – a separate `.R` or `.py` file, an appendix in the PDF, or as a `.Rmd` or similar format containing both content and code. This homework exercise is due by 11:59 p.m. ET on Wednesday, October 16, 2024; submit your materials in electronic format – via e-mail attachment – to Morrgan (`mth5492@psu.edu`) *and* to me (`zorn@psu.edu`). This exercise is worth 50 possible points.

[1] That is, you can generate a CDF of a Gumbel distributed variate from random draws from a $U_{0,1}$ distribution using this formula. This comment is a potentially valuable hint about completing this part of the exercise.