Tufts University
Department of Mathematics
Spring 2022

MA 166: Statistics

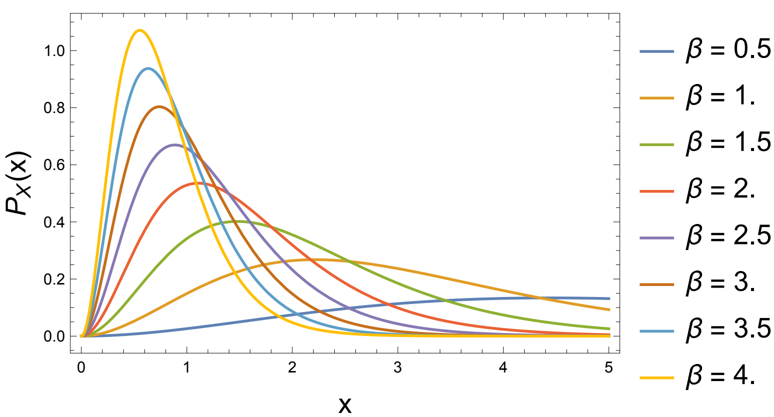
Solutions to Homework 1 (v1.0) [1]

Assigned Monday 24 January 2022
Due Monday 31 January 2022 at 11:59 pm EDT.

The first four problems of this homework assignment concern the random variable $X \geq 0$, which may be supposed to have the probability density function

$$f_X(x) = C(\beta)\frac{x^2}{1 + e^{\beta x}},$$

where β is a parameter and $C(\beta)$ is a normalization constant. This kind of random variable turns out to be important in physics for understanding the behavior of a certain class of fundamental particles called fermions. In that context, β is related to the inverse of the temperature of the fermions, so it is something that one might wish to measure experimentally. Some examples of this distribution for various values of β are illustrated below.



The following definite integrals may be useful for the problems in this assignment,

$$\begin{aligned}
\int_0^\infty dz\ \frac{z^2}{1+e^z} &= \frac{3}{2}\zeta(3) \\
\int_0^\infty dz\ \frac{z^3}{1+e^z} &= \frac{7\pi^4}{120} \\
\int_0^\infty dz\ \frac{z^4}{1+e^z} &= \frac{45}{2}\zeta(5),
\end{aligned}$$

where $\zeta(n) := \sum_{j=1}^{\infty} \frac{1}{j^n}$ is called the *Riemann zeta function*. You don't need to know anything about $\zeta(3)$ and $\zeta(5)$ except that they are constants, and you can simply write them as $\zeta(3)$ and $\zeta(5)$ most of the time, and not worry about their numerical values. In case you should need their numerical values, they are approximately $\zeta(3) = 1.2020569\ldots$ and $\zeta(5) = 1.0369277\ldots$.

1. **Find an expression for the normalization constant $C(\beta)$ so that $\int_0^\infty dx\, p_X(x)$ is equal to one. As the notation indicates, you should expect this normalization constant to depend on the parameter β. With this normalization constant, write the normalized distribution.**

   We demand that

$$
\begin{aligned}
1 =& \int_0^\infty dx\ f_X(x) \\
=& \int_0^\infty dx\ C(\beta) \frac{x^2}{1+e^{\beta x}} \\
=& C(\beta) \int_0^\infty \frac{dz}{\beta} \frac{(z/\beta)^2}{1+e^z} \\
=& \frac{C(\beta)}{\beta^3} \int_0^\infty dz\ \frac{z^2}{1+e^z} \\
=& \frac{3}{2}\zeta(3) \frac{C(\beta)}{\beta^3},
\end{aligned}
$$

   where we have used integration by substitution in the third step, with $x = z/\beta$ and hence $dx = dz/\beta$. It follows that

$$
\boxed{C(\beta) = \frac{2}{3\zeta(3)}\beta^3.}
$$

   The normalized density function is therefore

$$
\boxed{f_X(x) = \frac{2\beta^3}{3\zeta(3)} \frac{x^2}{1+e^{\beta x}},}
$$

2. **Now suppose that you have a list of experimental data, $X_j = x_j$ for $j = 1, \ldots, n$, which you expect to be distributed according to the given density function, but you do not know the value of the parameter β, and you would like to infer it from the data. Find an equation for the maximum likelihood estimate of β, call it β_e. (You will not be able to solve this equation in general for β_e. All you need for your answer is the equation itself.)**

   The likelihood function is

$$
L(\beta; \vec{k}) = \prod_{j=1}^n f_X(x_j) = \prod_{j=1}^n \frac{2\beta^3}{3\zeta(3)} \frac{x_j^2}{1+e^{\beta x_j}},
$$

so the log likelihood is

$$\ln L(\beta;\vec{k}) = \sum_{j=1}^{n} \ln f_X(x_j) = \sum_{j=1}^{n} \left[\ln\left(\frac{2}{3\zeta(3)}\right) + 3\ln\beta + 2\ln x_j - \ln\left(1+e^{\beta x_j}\right)\right]$$

Set the derivative with respect to β to zero to find the location of the maximum,

$$\frac{\partial}{\partial\beta} 0 = \ln L(\beta;\vec{k}) = \sum_{j=1}^{n}\left(\frac{3}{\beta} - \frac{x_j e^{\beta x_j}}{1+e^{\beta x_j}}\right) = \frac{3n}{\beta} - \sum_{j=1}^{n}\frac{x_j}{1+e^{-\beta x_j}}$$

This yields the following equation for the maximum likelihood estimate

$$\boxed{\frac{3}{\beta_e} = \frac{1}{n}\sum_{j=1}^{n}\frac{x_j}{1+e^{-\beta_e x_j}}}$$

3. **Repeat the last problem, but this time use method of moments to find an estimate for $\boldsymbol{\beta}$, call it $\boldsymbol{\beta_m}$. This time, you will be able to find a solution for $\boldsymbol{\beta_m}$, and you should provide that for your answer.**

   We have one parameter, so we will need one moment. We have

$$\begin{aligned} E(X) &= \int_0^\infty dx\; p_X(x)x \\ &= \int_0^\infty dx\; \frac{2\beta^3}{3\zeta(3)}\frac{x^3}{1+e^{\beta x}} \\ &= \frac{2\beta^3}{3\zeta(3)}\int_0^\infty \frac{dz}{\beta}\,\frac{(z/\beta)^3}{1+e^z} \\ &= \frac{2\beta^3}{3\zeta(3)\beta^4}\int_0^\infty dz\; \frac{z^3}{1+e^z} \\ &= \frac{7\pi^4}{180\zeta(3)\beta}. \end{aligned}$$

   So, according to the method of moments, we set

$$\frac{7\pi^4}{180\zeta(3)\beta} = \frac{1}{n}\sum_{j=1}^{n} x_j,$$

   and this yields the solution

$$\boxed{\beta_m = \frac{7\pi^4}{180\zeta(3)}\left(\frac{1}{n}\sum_{j=1}^{n} x_j\right)^{-1}}$$

4. **Suppose that you conduct $\boldsymbol{n}$ trials, and you miraculously find the same value $\boldsymbol{x_j = 1}$ all $\boldsymbol{n}$ times. You still won't be able to find an analytic solution**

to the equation you obtained for the maximum likelihood estimate β_e, but you should be able to reduce it to something very simple, so that you can solve what remains either graphically or using a root-finding calculator, or a software tool like *Mathematica*. Do so, and compare your numerical result to β_m, the method of moments estimate. Explain how you can make sense of the approximate values of your estimates from the plots of $p_X(x)$ given on the first page of this assignment.

All n terms in the sum in the equation for the maximum likelihood estimate are identical, so the sum divided by n is the same thing as the jth term in the sum, with $x_j = 1$. The equation for β_e therefore reduces to

$$\frac{3}{\beta_e} = \frac{1}{1 + e^{-\beta_e}}$$

Using a numerical root finder, it can be verified that the solution to the above is

$$\boxed{\beta_e = 3.1310197\ldots}$$

In similar fashion to the above, all n terms in the sum in the equation for the method of moments estimate are identical and equal to one, so the method of moments estimate is $\frac{7\pi^4}{180\zeta(3)}$, or

$$\boxed{\beta_m = 3.1513743\ldots}$$

As is evident from the above, the agreement between the two methods is rather remarkable for this problem. One can also examine the plots of $f_X(x)$ versus x for various values of β, provided on the first page of the assignment, to see why this result is sensible. The plot for $\beta = 3$ is the closest to the values of β estimated above, and that plot peaks very close to $x = 1$.