20/20

M525: HW1

Dominic Bair

September 26, 2022

1. Here we demonstrate our ability to load data and display it in different ways. We use the `vis_data.mat` file which contains 3 columns of data, W_i $i = 1, 2, 3$. We plot each column as a histogram, then plot each column against every other column as a scatter plot. See Figure (1).

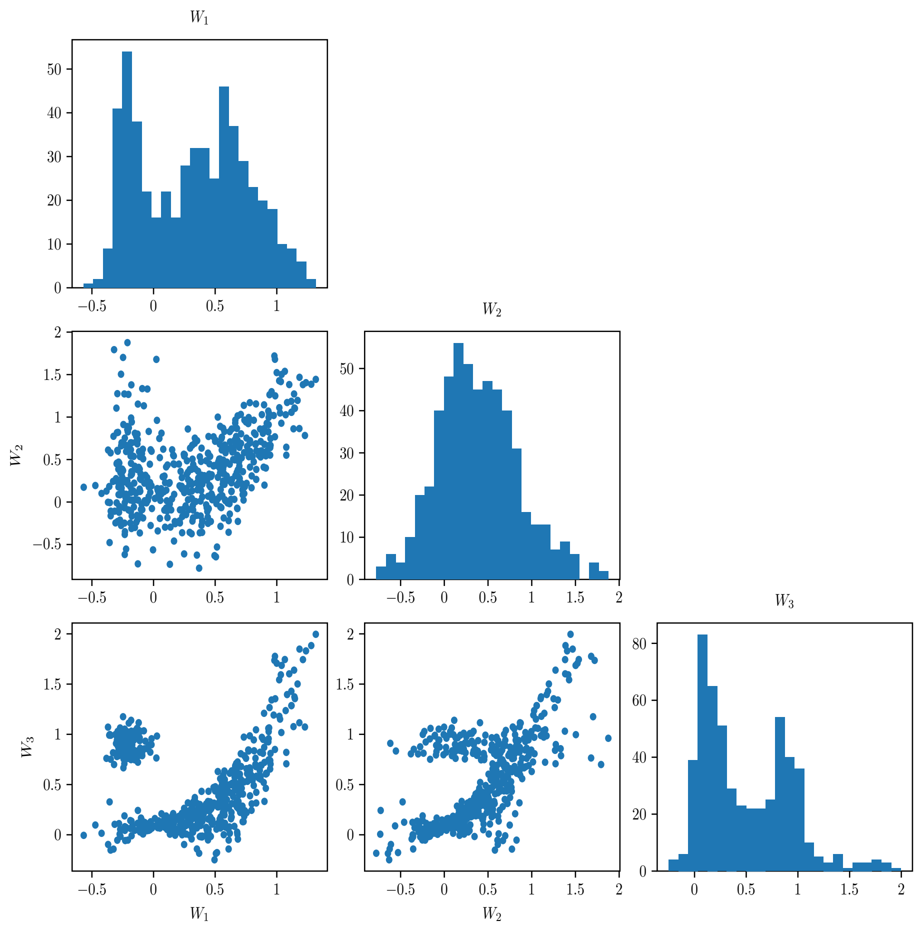

Figure 1: A visualization of `vis_data.mat`.

2. Here we generate and randomly sample categorical data. We choose seven categories, σ_i $i = 0, \ldots, 6$, with randomly generated probabilities of drawing a point from each category, π_i $i = 0, \ldots, 6$. We normalize our randomly generated probabilities so we can define a PMF. We then sample n random points from $(0, 1)$ and assign each point to a category based on the PMF; we accomplish this using MATLAB's `discretize()` function. See Figure (2).

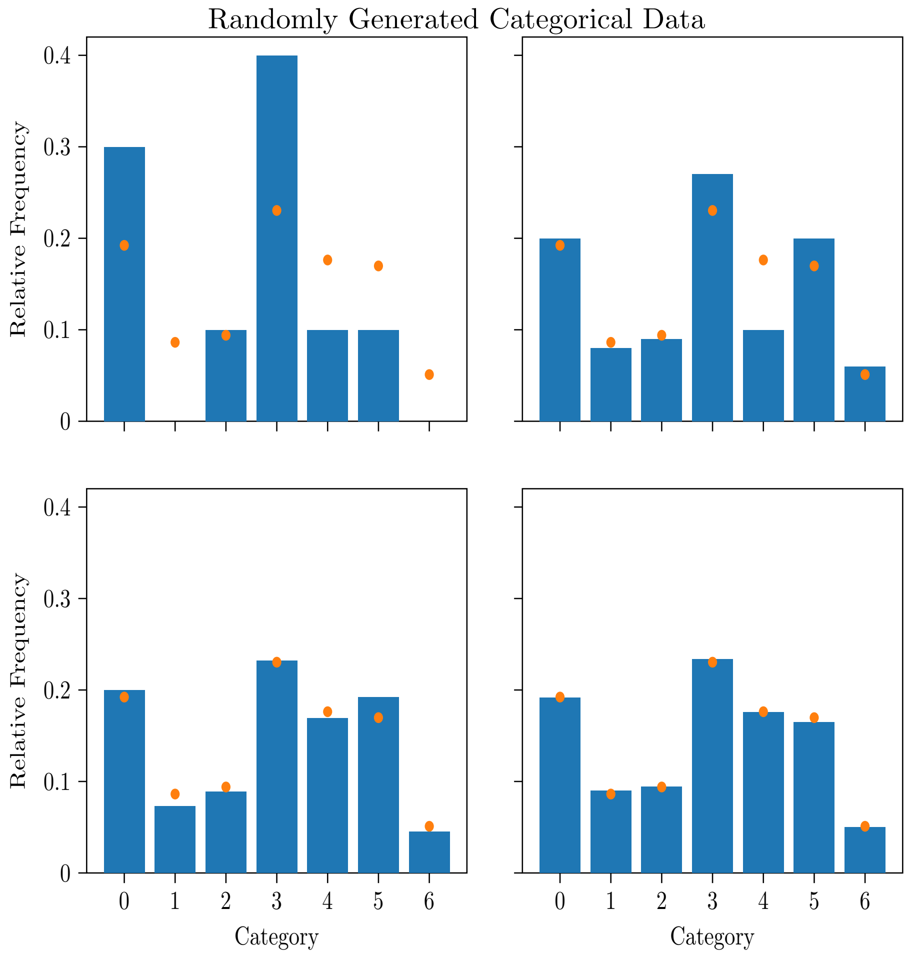

Figure 2: Bar plots of the relative frequencies of n randomly sampled categorical data points. There are seven possible categories a data point can take. Each orange dot represents the true probability of selecting an element from its corresponding category. (Top Left) $n = 10$. (Top Right) $n = 100$. (Bottom Left) $n = 1000$. (Bottom Right) $n = 10000$. Notice as the number of randomly sampled points increases, the relative frequency of a given category approaches the probability of a random point being a member of the category.

3. (a) First we show the Box–Muller algorithm produces random variables with the desired statistics. This notation is missleading

First, let $U, V \sim \text{Uniform}_{[0,1]}$, and let u and v be the values that U and V can take on respectively. The Box–Muller algorithm uses a map $f : U \times V \to X \times Y$ such that $f(u, v) = (\mu + \sigma\sqrt{-2\log u}\cos 2\pi v, \mu + \sigma\sqrt{-2\log u}\sin 2\pi v) = (x, y)$ to transform uniformly distributed random variables U, V to $X, Y \sim \text{Normal}(\mu, \sigma^2)$.

Proof. By the fundamental theorem of simulation we have

$$q(x, y) = \frac{p(u, v)}{|J_{(u,v)\to(x,y)}|}. \tag{1}$$

However,

$$\begin{aligned}
|J_{(u,v)\to(x,y)}| &= \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} \\
&= \frac{1}{u}\left(2\pi\sigma^2\cos^2 2\pi v + 2\pi\sigma^2\sin^2 2\pi v\right) \\
&= \frac{2\pi\sigma^2}{u}.
\end{aligned}$$

Also $p(u, v) = 1$ on its support, so by (1) we have

$$\begin{aligned}
q(x, y) &= \frac{u}{2\pi\sigma^2} \\
&= \frac{1}{2\pi\sigma^2}\exp\{\log u\} \\
&= \frac{1}{2\pi\sigma^2}\exp\left\{2\sigma^2\log u(\cos^2 2\pi v + \sin^2 2\pi v)\right\} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^2 \exp\left\{-\left(\frac{(\mu + \sigma\sqrt{-2\log u}\cos 2\pi v - \mu)^2}{2\sigma^2}\right)\right\} \\
&\times \exp\left\{-\left(\frac{(\mu + \sigma\sqrt{-2\log u}\sin 2\pi v - \mu)^2}{2\sigma^2}\right)\right\} \\
&= \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^2 \exp\left\{-\left(\frac{(x-\mu)^2}{2\sigma^2} + \frac{(y-\mu)^2}{2\sigma^2}\right)\right\} \\
&= \text{Normal}\left(x, y; \mu, \sigma^2\right).
\end{aligned}$$

We note that it is easier to solve this in reverse; that is starting with $\text{Normal}\left(x, y; \mu, \sigma^2\right)$ and working until you are left with (1); however, this is considered a faux pas as we would be assuming our result. ■

(b) We create a MATLAB function that performs the Box–Muller algorithm. We then plot 10,000 points using the function and plot the results in Figure 3.

(c) Here we show our results from our Box–Muller function.

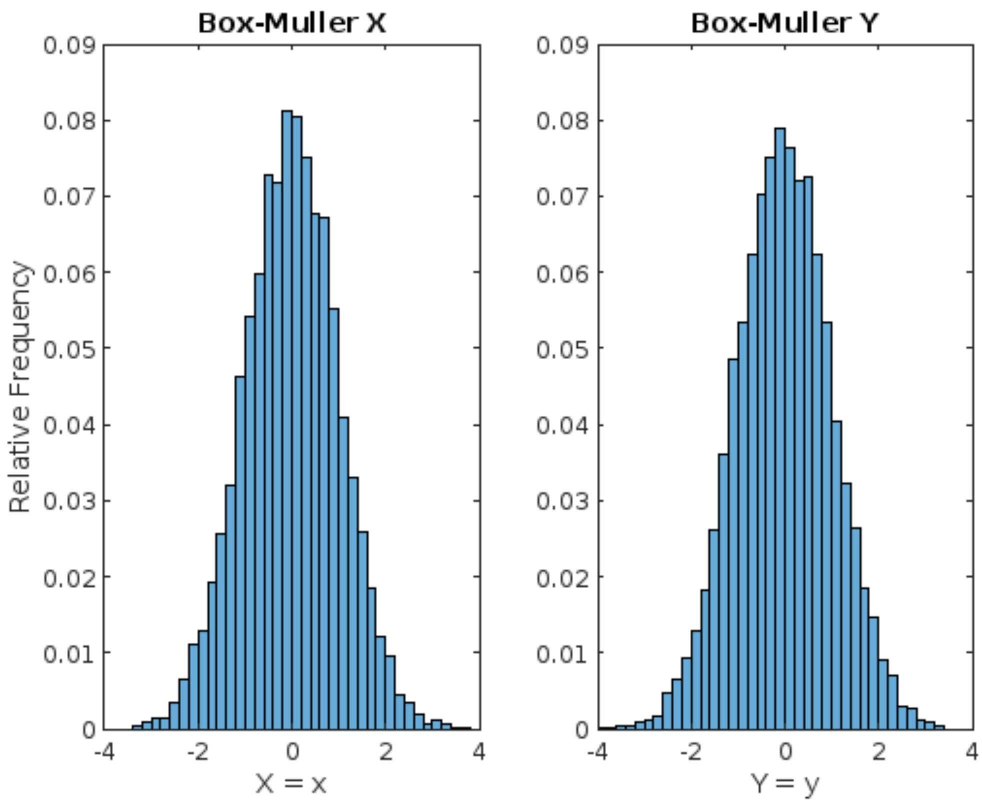

Figure 3: Both outputs of the Box–Muller algorithm developed in part (b). This simulation samples 10,000 points from a uniform distribution, then applies the Box–Muller algorithm with $\mu = 0$, and $\sigma = 1$. Both plots clearly have a mean centered at 0, and inflection points at ± 1. This result is as expected from our calculation in (a).

4. Here we analyse a model of molecule excitation and detecting photons emitted by the excited molecules. Our recorded measurement w_i for a photon pulse encodes the time d_i that a molecule remains excited, and an error, or noise, r_i caused by our detector. Thus our final measured signal is, $w_i = d_i + r_i$. We assume $D \sim \text{Exponential}(\lambda)$ and $R \sim \text{Normal}(0, v)$.

   (a) First, we analytically derive the "best" choice of our model parameters $\hat{\lambda}$ and $\hat{v}$ using the principle of maximum likelihood.

   First we examine our model for the random variable D. We assume each measurement is independent so we have

$$\begin{aligned}
\mathbb{P}(D = d_{1:N}) &= \prod_{i=1}^{N} \lambda \exp\{-\lambda d_i\} \\
&= \lambda^N \exp\left\{-\lambda \sum_{i=1}^{N} d_i\right\}. \\
\implies \log \mathbb{P}(D = d_{1:N}) &= N \log \lambda - \lambda \sum_{i=1}^{N} d_i. \\
\implies \frac{d}{d\lambda}(\log \mathbb{P}(D = d_{1:N})) &= \frac{N}{\lambda} - \sum_{i=1}^{N} d_i \overset{\text{set}}{=} 0. \\
\implies \hat{\lambda} &= \frac{N}{\sum_{i=1}^{N} d_i}.
\end{aligned}$$

These are irrelevant

   Next we examine our model for the random variable R. Again, we assume each measurement is independent so we have

$$\begin{aligned}
\mathbb{P}(R = r_{1:N}) &= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi v}} \exp\left\{\frac{-r_i^2}{2v}\right\} \\
&= \left(\frac{1}{\sqrt{2\pi v}}\right)^N \exp\left\{\frac{-\sum_{i=1}^{N} r_i^2}{2v}\right\}. \\
\implies \log \mathbb{P}(R = r_{1:N}) &= N \log \frac{1}{\sqrt{2\pi v}} + \left(\frac{\sum_{i=1}^{N} r_i^2}{2v}\right). \\
\implies \frac{d}{dv}(\mathbb{P}(R = r_{1:N})) &= N\sqrt{2\pi v}\frac{1}{\sqrt{2\pi}}\left(\frac{-1}{2}\right) v^{\frac{-3}{2}} + \frac{v^{-2}}{2} \sum_{i=1}^{N} r_i^2 \\
&= \frac{-N}{v} + \frac{\sum_{i=1}^{N} r_i^2}{v^2} \overset{\text{set}}{=} 0. \\
\implies \hat{v} &= \frac{\sum_{i=1}^{N} r_i^2}{N}.
\end{aligned}$$

   Thus, we have found the "best" parameters for our models of the distributions of D and R. Applying our calibration data to this, we find $\hat{v} = 1.6581$.

Add units

(b) We now derive the probability density formula for the random variable W.

To use the fundamental theorem of simulation we must have an invertible map, so we introduce a new random variable, $Z = D$. Then we have the map $f : \mathcal{R} \times \mathcal{D} \to \mathcal{W} \times \mathcal{Z}$ with $f(r, d) = (r + d, d) = (w, z)$. Using the fundamental theorem of simulation we find

$$p(w, z) = \lambda \exp\{-\lambda d\} \frac{1}{\sqrt{2\pi v}} \exp\left\{\frac{-r^2}{2v}\right\}.$$

We then marginalize this to find $p(w)$.

$$\begin{aligned} p(w) &= \int_0^\infty \lambda \exp\{-\lambda z\} \frac{1}{\sqrt{2\pi v}} \exp\left\{\frac{-(w-z)^2}{2v}\right\} \mathrm{d}z \\ &= \frac{\lambda}{2} \exp\left\{\lambda^2 v^2 - \lambda w\right\} \left(1 - \operatorname{erf}\left(\frac{\lambda v - w}{\sqrt{2v}}\right)\right). \end{aligned}$$

(c) Now we apply the principle of maximum likelihood to this density to find our optimal $\hat{\lambda}$. We note this $\hat{\lambda}$ will be different from our above problem as this is a new distribution. We use our $\hat{v}$ from above as we calibrate our instrument first.

$$\begin{aligned} \mathbb{P}(W = w_{1:N}) &= \prod_{i=1}^{N} \frac{\lambda}{2} \exp\left\{\lambda^2 \hat{v}^2 - \lambda w_i\right\} \left(1 - \operatorname{erf}\left(\frac{\lambda \hat{v} - w_i}{\sqrt{2\hat{v}}}\right)\right). \\ \implies -\log\left(\mathbb{P}(W = w_{1:N})\right) &= N \log\left(\frac{\lambda}{2}\right) + \sum_{i=1}^{N} \left(\frac{\lambda^2 \hat{v}}{2} - \lambda w_i\right) \\ &\quad + \sum_{i=1}^{N} \log\left(1 - \operatorname{erf}\left(\frac{\lambda \hat{v} - w_i}{\sqrt{2\hat{v}}}\right)\right). \end{aligned}$$

We now have the negative log–likelihood of our data, so we find $\hat{\lambda}$ using `fminsearch()` in MATLAB. Applying our experimental data, we find $\hat{\lambda} = 0.8086$. See Figure 4.

Units

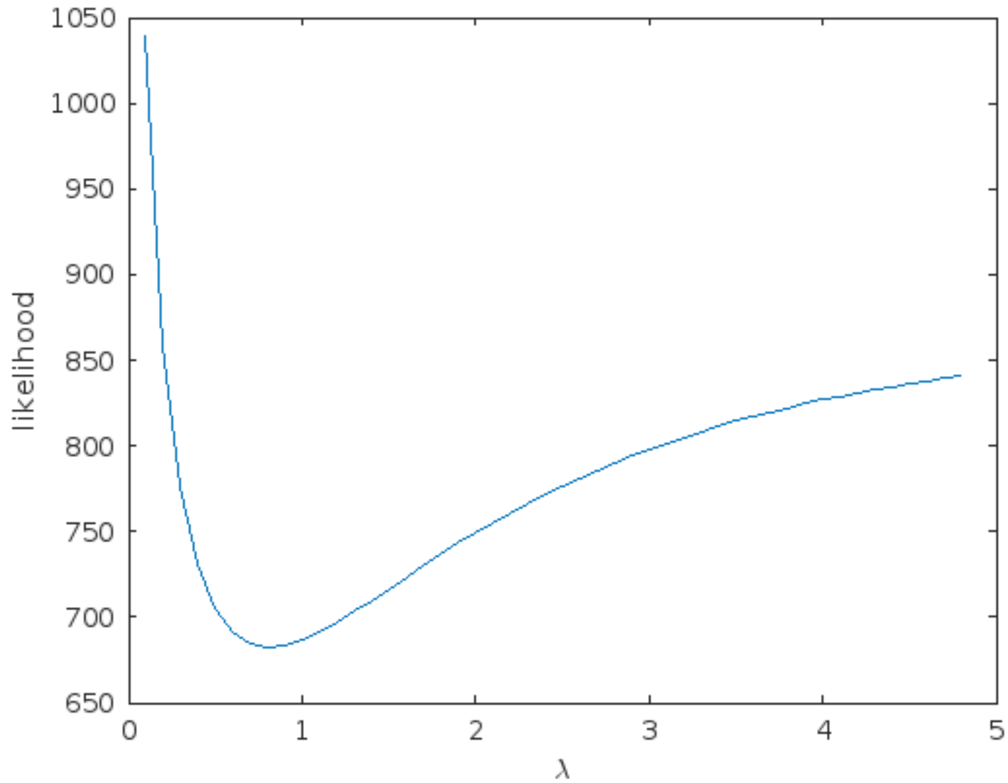

Figure 4: A plot of our negative log–likelihood for our data. We see visually the minimizer, $\hat{\lambda}$, should be near 0.8.