Assignment 4 - Experimentation

William M. Temple

May 4, 2018

1. Design

Introduction and Procedure

In this experiment, I design and implement an apparatus to test how experiment participants' recover numerical information from graphical visualizations of data. In particular I will compare two visualization strategies for bivariate Gaussian distributions with independent dimensions ($\rho\sigma_x\sigma_y = \rho\sigma_y\sigma_x = 0$). In other words, the bivariate Gaussian distributions have the following form:

$$N((\mu_x, \mu_y), (\sigma_x, \sigma_y)$$

$$V_{xy} = \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix}$$

I consider two visualization strategies:

- A single image-plot utilizing a color-map from white to black to show probability density (Figure 1).

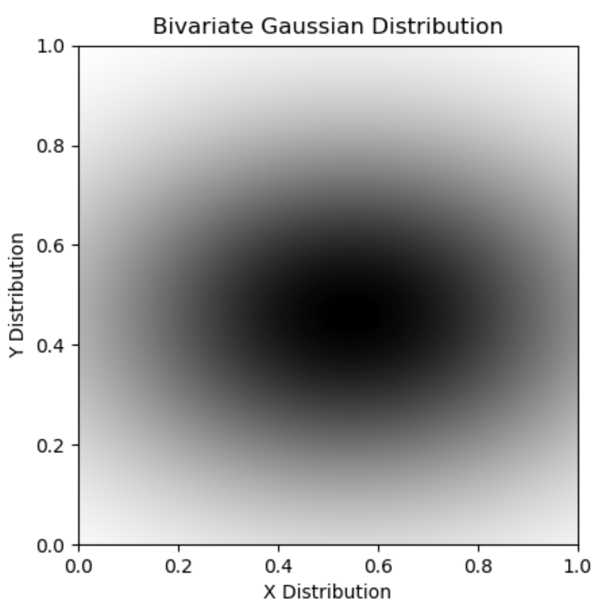


Figure 1: 2-Dimensional image plot of Bivariate Gaussian Density

- Two separate views each showing the X and Y components of the graphs only (Figure 2)

The experimental procedure is simple:

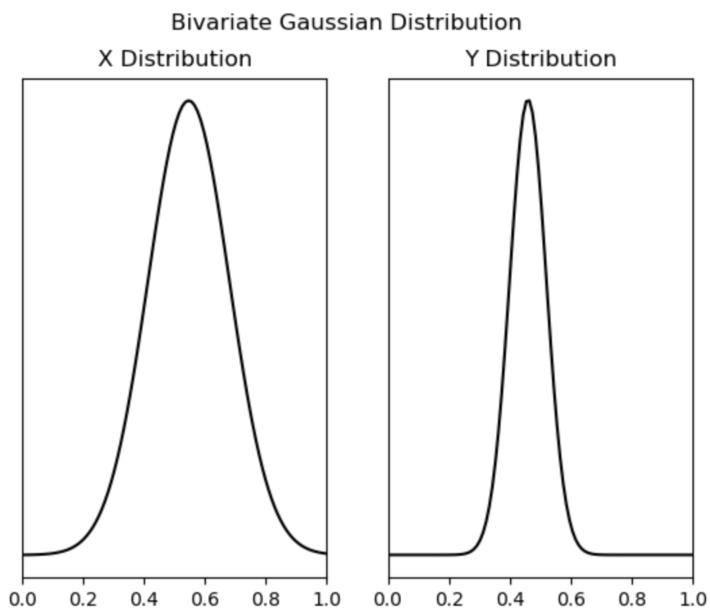


Figure 2: Two coordinated views of X and Y data only

1. I will present the participant with some examples of both of these plots, one at a time, alternating between image-plot and X/Y subplots.
2. I will then ask the participant to estimate the mean (μ) and standard-deviation (σ) for both the X and Y distributions.
3. I will record the estimated mean and standard-deviation for later analysis.

The experiment will conclude after six iterations of the above procedure per participant.

Experiment Overview

In this experiment, I wish to explore whether or not coordinated X and Y subplots make it easier for a participant to estimate the mean and standard-deviation of the X and Y components of a bivariate Gaussian distribution compared to an image map. More specifically, are participants' estimations of mean and standard-deviation for the X and Y components of a bivariate Gaussian distribution closer to the actual mean and standard deviation when the participant is presented with an image map of the joint distribution or when a participant is presented with two corresponding X and Y subplots? I hypothesize that the participants will make more accurate estimations when using the two separate X and Y subplots, as the individual components may be analyzed independently with less visual interference from the other competing dimension in the image-plot.

The independent variable is the kind of plot shown to the participant, either an image-plot or X/Y subplots. The dependent variables are the distances between estimated and actual values of μ and σ. I will control other variables to ensure consistency in other aspects of data presentation:

1. Actual probability densities will remain unlabeled, and I will scale the means of each distribution in the X/Y subplots so that they match, as the image-plot combines the probability densities in the 2D image space. Removing axis-labels ensures that noticeable differences in the maximum probability density for the X/Y subplots do not affect the estimated mean or standard-deviation. Practically, this

means that the image-plot has no color-bar and the y-axes of the X/Y subplots do not have axis ticks or labels, nor do they share a y-axis.

2. In order to avoid any suggestions made by the use of color, all plots will use only greyscale.

3. I label the x-axis and y-axis of the image-plot in the same way that I label the X and Y subplots, respectively, and the overall graphs are titled identically, to avoid any auxiliary effects of word choice.

Due to the small study population I anticipate ($5 < N < 10$), any statistical significance is unlikely to emerge. However, I will report the participants' average estimation errors and compute a confidence interval for the estimation error values for each visualization type. I will explain, qualitatively, how the data do or do not support my hypothesis and whether or not they justify further experimentation.

2. Apparatus

For this experiment, I used an online survey tool (Google Forms) to simplify data collection. I wrote a Python program (using matplotlib, numpy, and scipy) to generate figures (like Figure 1 and 2), and embedded them into the online survey. Each question asks the participant to estimate the mean and standard-deviation of both the X and Y components of the distribution. The questions only accept numbers within the interval $[0, 1]$. The full survey text is embedded below (Appendix A).

3. Study and Results

I recruited six participants from among friends, former colleagues, and members of the Laboratory for Playful Computation. No knowledge of statistics was required to participate. I averaged the distance between the participants' responses and the actual answers, and compared the average error when participants were presented with the coordinated plots to the same when participants were presented with the image plot.

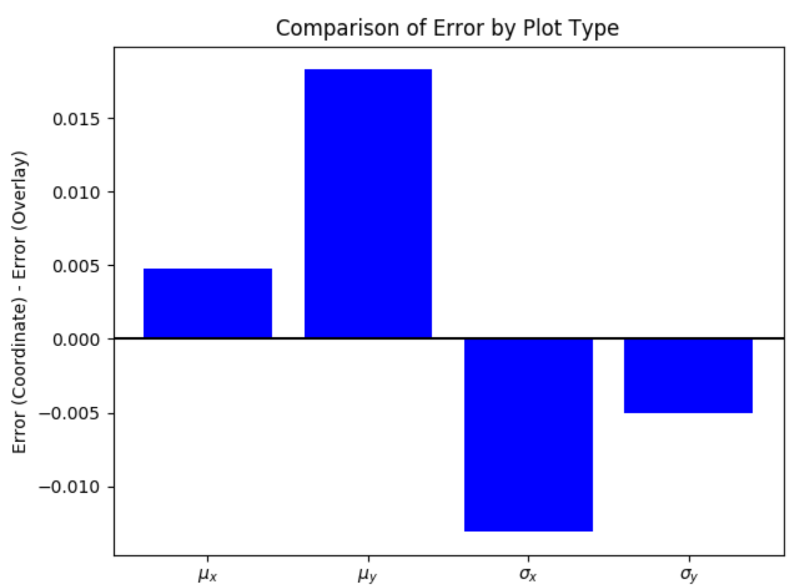


Figure 3: Average error difference in plot types

The graph shows the average difference (Figure 3). Values below zero indicate that the image plots performed better and values above zero indicate that the coordinated plots performed better.

The data immediately restrict my hypothesis: in the small sample of participants who generated this data, the image plot was associated with more accurate estimations for the X distribution standard-deviation (σ_x) and Y standard-deviation (σ_y).

4. Analysis and Discussion

The first, most obvious conclusion to draw from these data are that far more data are required to make any accurate statements regarding my hypothesis. With only six participants, each making estimations regarding three of each plot type, the total number of samples is only 18 for each plot type. As such, I've chosen to focus on a qualitative discussion of the results and offer some suggestions about what effects could possibly influence these results.

Estimating Standard-Deviation

I found that participants consider the standard-deviation difficult to estimate. While recorded estimates for mean-values exhibited little variance ($\sigma < 0.05$), estimates for standard-deviation exhibited much higher variance ($\sigma > 0.1$ for some questions). One participant even remarked that the questions are "evil" and resorted to using a makeshift straightedge (a *Post-it* Note) to align the mean and estimate the standard deviation from there.

I acknowledge that this task is probably too difficult to make accurate predictions about how the graph type affects one's ability to estimate standard-deviation (as estimations have high variance *regardless* of plot type). It may have been useful or instructive to provide some examples which show the standard-deviation of some sample distributions. I chose not to do this in this experiment because I was worried that such instruction for each chart type would make the estimations too accurate. However, I think that estimations which are too accurate would be more useful than these, which are too inaccurate and too highly varied.

Estimating the Mean

Participants were more easily able (though certainly not *significantly*) to estimate an accurate mean for distributions when the distributions were presented in two separate subplots. One participant remarked that this task was easy in this case, because the mean of the distribution is obviously the single modal peak. It may be more interesting to consider these tasks with multi-modal distributions, though this risks elevating the difficulty beyond reasonable levels.

In summary, I do not confirm my hypothesis and, broadly speaking, require much more data to conduct a convincing analysis. Due to the rote nature of task performance for this experiment, it may be a good candidate for Mechanical Turk or other similar services.

Appendix A

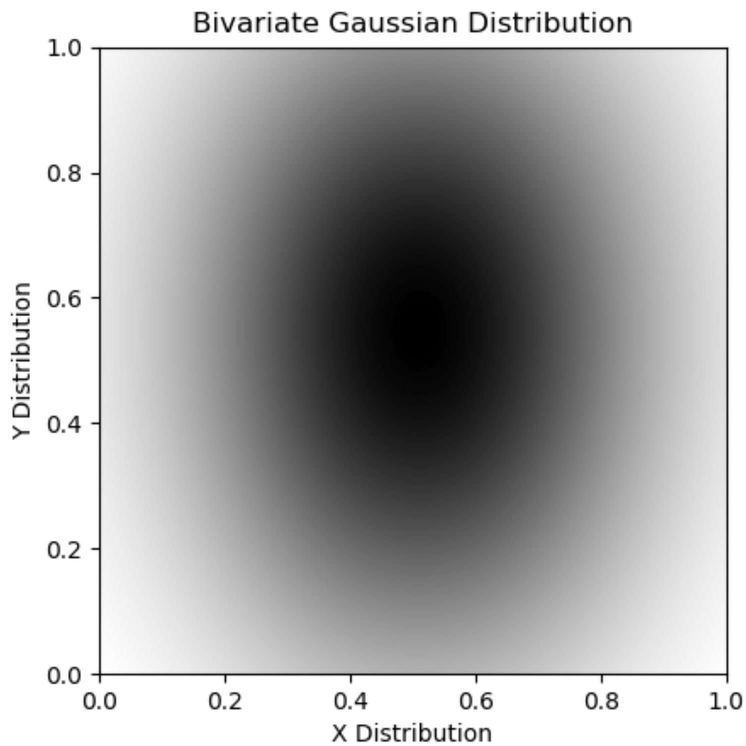
Normal Distribution Survey

In this survey, we are seeking to understand people's ability to estimate the mean and standard-deviation of normally-distributed data. In the following questions, you will be shown some graphs that describe 2D Gaussian distributions. You will then be asked to

* Required

Question 1/6

Answer the following questions about the graph:



1. **Estimate the mean of the X distribution:** *

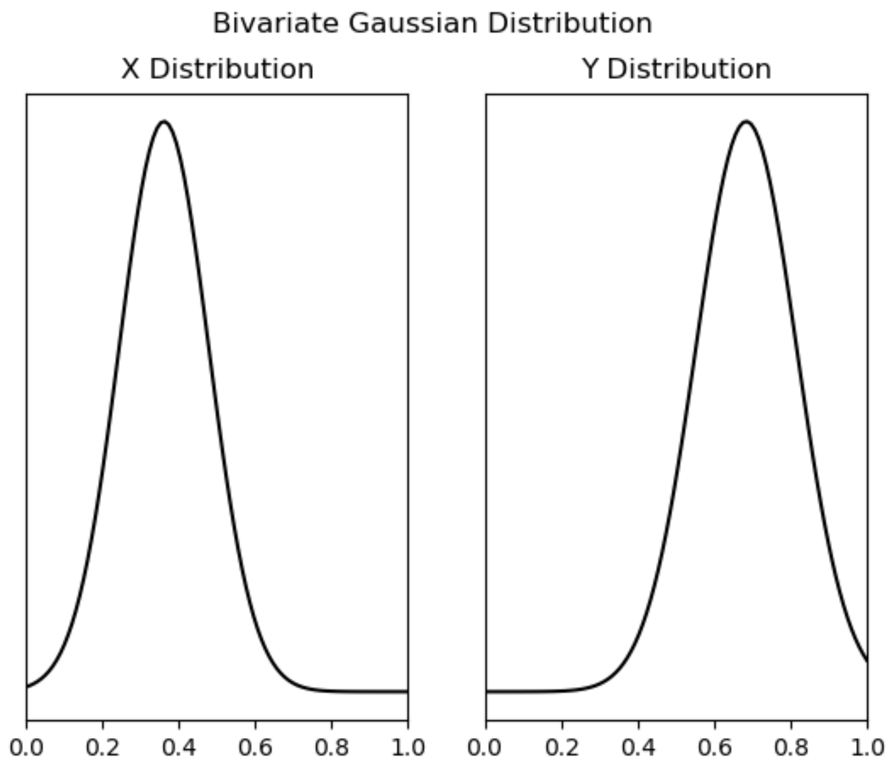
2. **Estimate the standard-deviation of the X distribution:** *

3. **Estimate the mean of the Y distribution:** *

4. **Estimate the standard-deviation of the Y distribution:** *

Question 2/6

Answer the following questions about the graph:



5. **Estimate the mean of the X distribution:** *

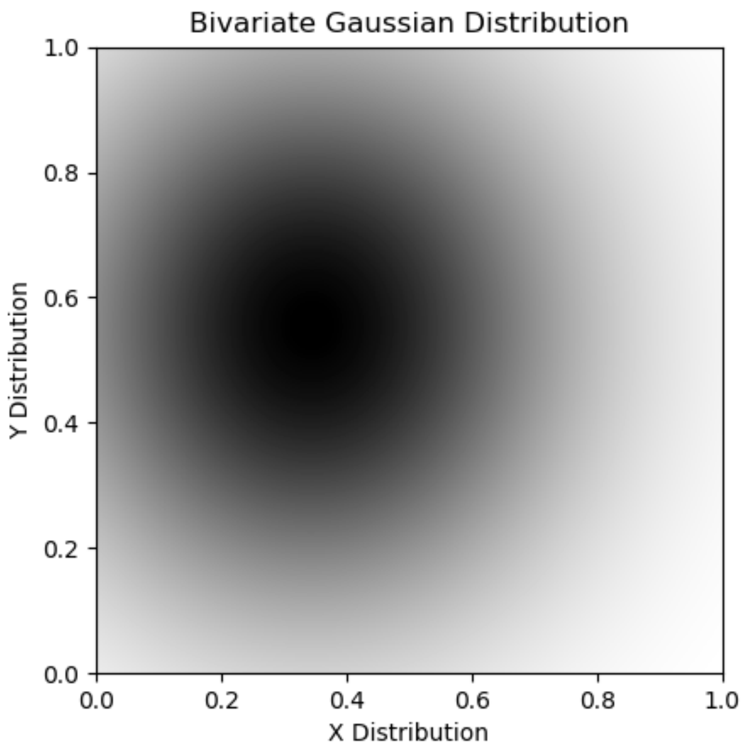
6. **Estimate the standard-deviation of the X distribution:** *

7. **Estimate the mean of the Y distribution:** *

8. **Estimate the standard-deviation of the Y distribution:** *

Question 3/6

Answer the following questions about the graph:



9. **Estimate the mean of the X distribution:** *

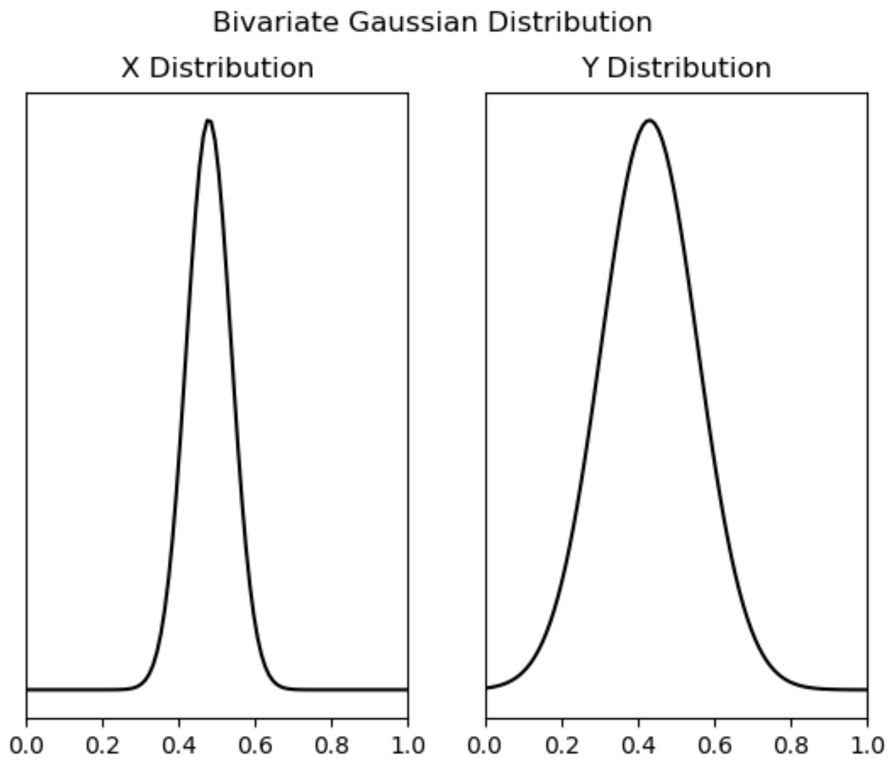
10. **Estimate the standard-deviation of the X distribution:** *

11. **Estimate the mean of the Y distribution:** *

12. **Estimate the standard-deviation of the Y distribution:** *

Question 4/6

Answer the following questions about the graph:



13. **Estimate the mean of the X distribution:** *

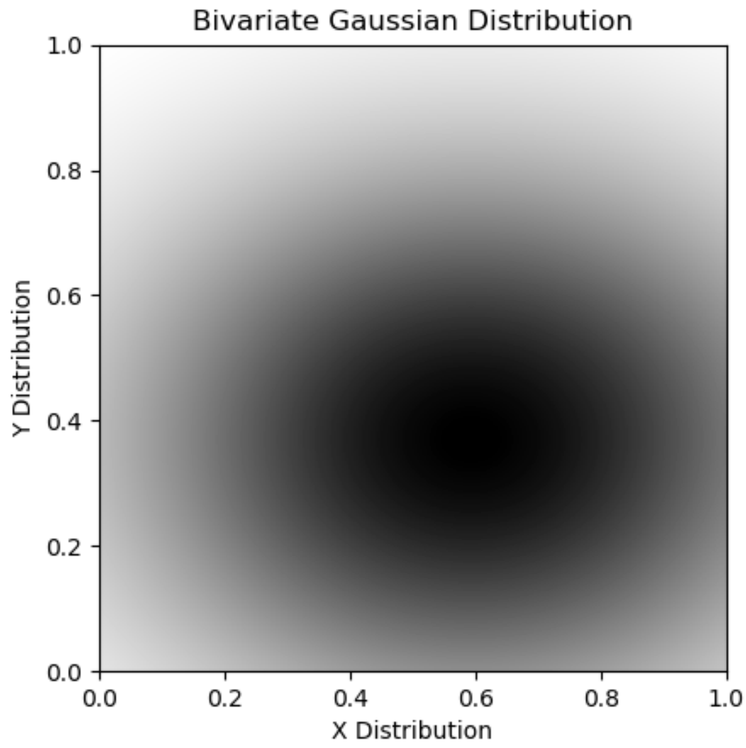
14. **Estimate the standard-deviation of the X distribution:** *

15. **Estimate the mean of the Y distribution:** *

16. **Estimate the standard-deviation of the Y distribution:** *

Question 5/6

Answer the following questions about the graph:



17. **Estimate the mean of the X distribution:** *

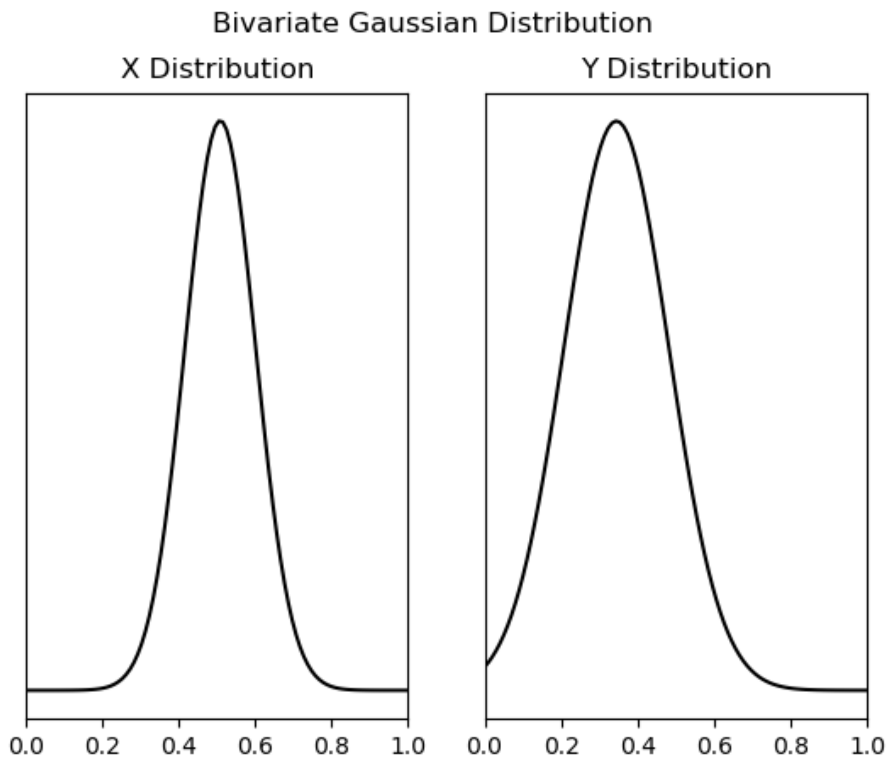
18. **Estimate the standard-deviation of the X distribution:** *

19. **Estimate the mean of the Y distribution:** *

20. **Estimate the standard-deviation of the Y distribution:** *

Question 6/6

Answer the following questions about the graph:



21. **Estimate the mean of the X distribution:** *

22. **Estimate the standard-deviation of the X distribution:** *

23. **Estimate the mean of the Y distribution:** *

24. **Estimate the standard-deviation of the Y distribution:** *