Behavioral Data Science Homework 1
Kelly Modi, Dev Dwivedi
31 Jan 2025

Problem 1: True/False Qs

1. False – As discussed in class, even papers published in peer-reviewed journals showed low reproducibility rate
2. True
3. False – Continuous random variables require an *uncountable* support and probability *density* function
4. True
5. False – Representation learning focuses on automatically extracting raw data instead of manually extracting it in handcrafted feature engineering
6. True
7. True
8. False – Variables and operations are represented as nodes and edges represent the flow of data between operations in a DAG.

Problem 2: Replication Crisis

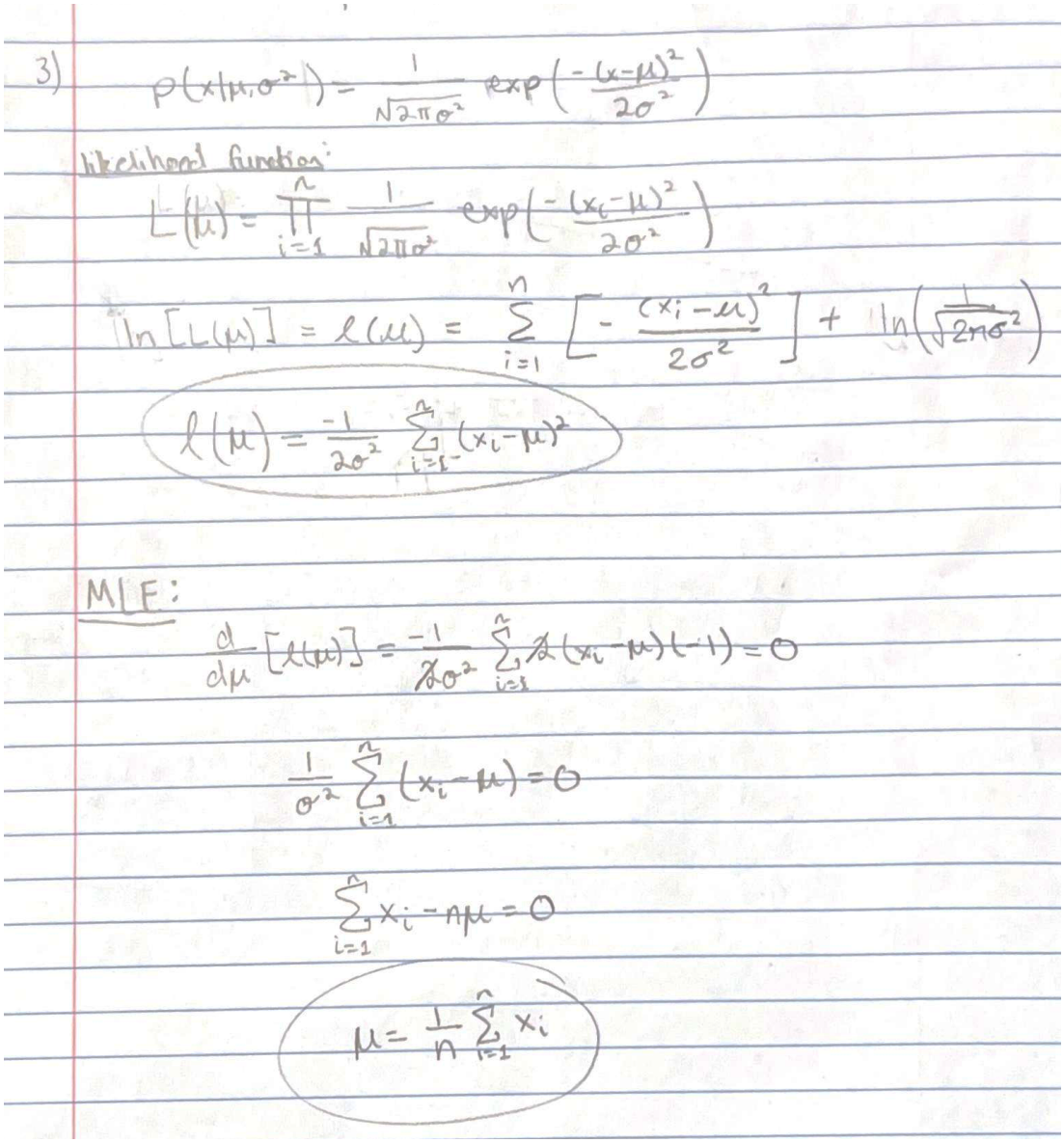
How can theory building and statistical methodology synergize to support strong and reproducible psychological science?

Theory building involves deriving testable claims that explain how and why a phenomenon or pattern occurs. While it provides a useful framework for understanding concepts, the replication crisis has shown that theoretical models are not always reproducible. Statistical methodology focuses on methods for collecting and analyzing data to draw conclusions. Integrating theory building with statistical methodology is crucial for developing a strong and reproducible psychological science.

New statistical methods allow theories to be tested on a larger and more complex scale, leading to more generalizable research findings. For example, statistical methodologies have evolved from parametric models to machine learning techniques to handle high-dimensional data. Additionally, computational graphs, such as Directed Acyclic Graphs, help refine theoretical claims.

To address the replication crisis, there has been a shift from a data modeling culture to an algorithm modeling culture. This transition prioritizes models that focus on prediction rather than explanation. By leveraging new statistical tools and emphasizing predictive modeling, psychological science can become more robust and replicable.

Problem 3: Maximum Likelihood Estimation



3) $p(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$

likelihood function:

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x_i-\mu)^2}{2\sigma^2}\right)$$

$$\ln[L(\mu)] = \ell(\mu) = \sum_{i=1}^{n} \left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] + \ln\left(\sqrt{\frac{1}{2\pi\sigma^2}}\right)$$

$$\ell(\mu) = \frac{-1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2$$

MLE:

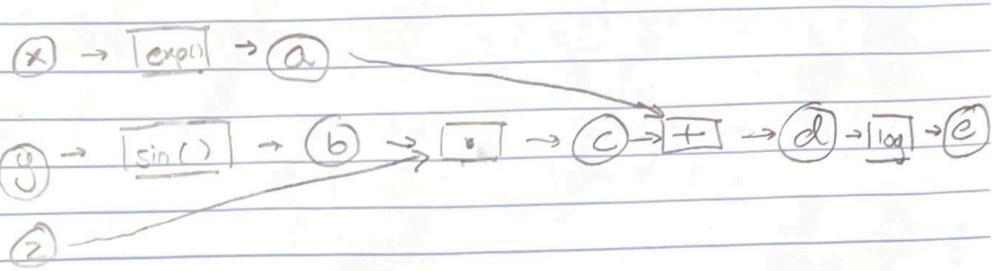
$$\frac{d}{d\mu}[\ell(\mu)] = \frac{-1}{2\sigma^2} \sum_{i=1}^{n} 2(x_i-\mu)(-1) = 0$$

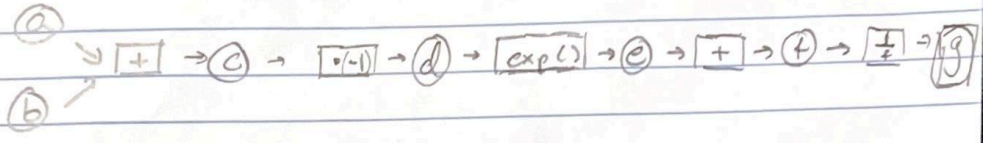
$$\frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i-\mu) = 0$$

$$\sum_{i=1}^{n} x_i - n\mu = 0$$

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

4) Part 2:

1. 

2. 

7) $y = B_0 + B_1 x_1 + B_2 x_2 \longrightarrow y = B_0 + B_1 x_1 + B_2 x_2 + B_3 x_1 x_2$

x_1	x_2	$x_1 x_2$	y	
0	0	0	0	$= B_0$
0	1	1	1	$= B_0 + B_2$
1	1	1	0	$= B_0 + B_1 + B_2 + B_3$
1	0	1	1	$= B_0 + B_1$

$y = x_1 + x_2 - 2(x_1 x_2)$

$B_0 = 0$

$B_1 = 1$

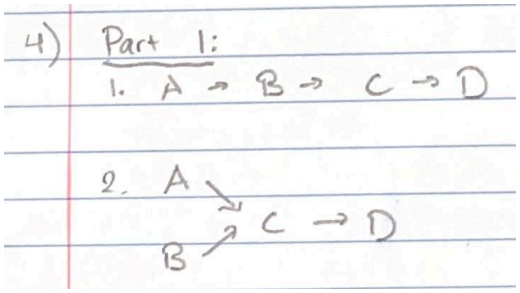
$B_2 = 1$

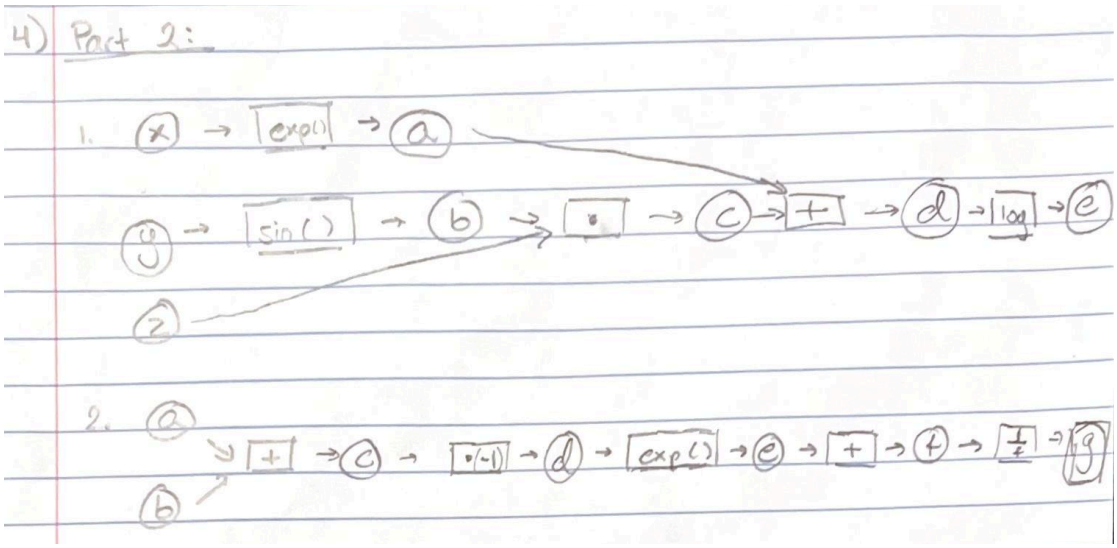
$B_3 = -2$

$B_0 = 0 \qquad 1 = 0 + B_2 \rightarrow B_2 = 1 \qquad 1 = 0 + B_1 \rightarrow B_1 = 1$

$0 = 0 + 1 + 1 + B_3 \rightarrow B_3 = -2$

Problem 4: Graphical Models and Computational Graphs





Problem 5: Git and GitHub

Part 1: Using Git and Creating Team Github

https://github.com/dev-dwivedi/homework-1.git

Part 2: Merge Conflict Resolution Attempts:

```
C:\Users\devdw\Downloads\Behavioral Data Science\Homework 1>git pull origin Dev-Dwivedi --rebase
From https://github.com/dev-dwivedi/homework-1
 * branch            Dev-Dwivedi -> FETCH_HEAD
Auto-merging README.MD
CONFLICT (content): Merge conflict in README.MD
error: could not apply a1cffb9... Editted README in Dev Branch
hint: Resolve all conflicts manually, mark them as resolved with
hint: "git add/rm <conflicted_files>", then run "git rebase --continue".
hint: You can instead skip this commit: run "git rebase --skip".
hint: To abort and get back to the state before "git rebase", run "git rebase --abort".
hint: Disable this message with "git config advice.mergeConflict false"
Could not apply a1cffb9... Editted README in Dev Branch

C:\Users\devdw\Downloads\Behavioral Data Science\Homework 1>
```

```
Editted README in Dev Branch

# Conflicts:
#       README.MD

# Please enter the commit message for your changes. Lines starting
# with '#' will be ignored, and an empty message aborts the commit.
#
# interactive rebase in progress; onto 3605921
# Last command done (1 command done):
#    pick a1cffb9 Editted README in Dev Branch
# No commands remaining.
# You are currently rebasing branch 'Dev-Dwivedi' on '3605921'.
#
# Changes to be committed:
#       modified:   README.MD
#
~
~
~
~
~
~
~
~
~
~
~
~
.git/COMMIT_EDITMSG [unix] (12:26 30/01/2025)
"~/Downloads/Behavioral Data Science/Homework 1/.git/COMMIT_EDITMSG" [unix] 17L, 469B
```

```
C:\Users\devdw\Downloads\Behavioral Data Science\Homework 1>git rebase --continue
[detached HEAD b681918] Editted README in Dev Branch
 1 file changed, 1 insertion(+), 2 deletions(-)
Successfully rebased and updated refs/heads/Dev-Dwivedi.

C:\Users\devdw\Downloads\Behavioral Data Science\Homework 1>
```

```
C:\Users\devdw\Downloads\Behavioral Data Science\Homework 1>git push origin Dev-Dwivedi
Enumerating objects: 5, done.
Counting objects: 100% (5/5), done.
Delta compression using up to 16 threads
Compressing objects: 100% (3/3), done.
Writing objects: 100% (3/3), 394 bytes | 394.00 KiB/s, done.
Total 3 (delta 1), reused 0 (delta 0), pack-reused 0 (from 0)
remote: Resolving deltas: 100% (1/1), completed with 1 local object.
To https://github.com/dev-dwivedi/homework-1.git
   3605921..b681918  Dev-Dwivedi -> Dev-Dwivedi

C:\Users\devdw\Downloads\Behavioral Data Science\Homework 1>
```

```
C:\Users\modik\Downloads\Behavioral Data Science\homework-1>git pull origin Kelly-Modi
From https://github.com/dev-dwivedi/homework-1
 * branch            Kelly-Modi -> FETCH_HEAD
Auto-merging README.MD
CONFLICT (content): Merge conflict in README.MD
Automatic merge failed; fix conflicts and then commit the result.

C:\Users\modik\Downloads\Behavioral Data Science\homework-1>git status
On branch main
Your branch is up to date with 'origin/main'.

All conflicts fixed but you are still merging.
  (use "git commit" to conclude merge)

Changes to be committed:
        modified:   README.MD
        new file:   monte_carlo.ipynb


C:\Users\modik\Downloads\Behavioral Data Science\homework-1>git pull origin Kelly-Modi
error: You have not concluded your merge (MERGE_HEAD exists).
hint: Please, commit your changes before merging.
fatal: Exiting because of unfinished merge.

C:\Users\modik\Downloads\Behavioral Data Science\homework-1>git commit -m "resolving merge conflict"
[main f115c36] resolving merge conflict
```

Merge Conflict Resolved

Part 3: Defining Git Commands

- Git restore:
    - This command can be used to remove uncommitted changes in the working directory. For example, if a text file was modified but was not staged in the staging area, git restore [file] would restore it back to the last committed version of that file.
- Git checkout:
    - This command is used to switch to a branch. For example: git checkout main.
- Git reset:
    - This command removes the specified file from the staging area, undoing any changes made since the last commit, but does not delete the changes made to a file. For example, I would use the git reset [file] to remove a file in the staging area.

- Git revert:
    - This command creates a new commit that undoes changes made in the indicated commit. It does not delete the indicated commit but it creates a new commit that reverts the changes made in that commit. To do this, I would use the command git revert [commit].

Problem 6: Python and NumPy

In Visual Studio and GitHub Homework File

Problem 7: Linear Model for XOR