Ryan Rowe & Chase Grajeda
COGS 4270-01
2024-02-12

Assignment 1 (Non-Programming Questions)

[Q1] True-False Questions [4 pts]
Mark all statements which are **FALSE**.

1. A random variable is discrete if its support is countable and there exists an associated probability density function (pdf).
2. Probability mass functions have a lower bound of 0 and an upper bound of 1.
3. The set of all possible realizations of a random variable is called its probability density.
4. The expected value of a discrete random variable is always part of its support, that is, $\mathbb{E}[X] \in R_X$.
5. Continuous random variables are functions which map points from the sample space to the real numbers.
6. We can formulate most parametric Bayesian models as a generative process, by which we first sample from the likelihood and then use the synthetic data point to sample from the prior.
7. The Bayesian posterior $p(\theta \mid y)$ for continuous parameter vectors $\theta \in \mathbb{R}^D$ is just another density function. That means, its integral $\int p(\theta \mid Y = y)\, d\theta \neq 1$ for some y.
8. Each realization of a continuous random variable has a probability of 0.

1. **False**. [W03 S08] Definition 2.2 (discrete random variable): "A random variable X is discrete if its support R_X is a countable set AND there is a function $p : R \to [0, 1]$ called the probability mass function (pmf) of X."
2. True. [W03 S08].
3. **False**. The set of all possible realizations of a random variable is actually called its support.
4. **False**. The mean of X ($\mathbb{E}[X]$) doesn't have to be a value that X can be (R_X). Example: If X had equal probability chance of being 1 or 2, then $\mathbb{E}[X] = 1.5$, which is not in $R_X = \{1, 2\}$.
5. True.
6. True.
7. **False**.
8. True. [W03 S17].

Assignment 1 (Non-Programming Questions)

[Q2] Git and GitHub [8 pts]

1. Create a public GitHub repository, create and add a team logo to the `README` file, along with some basic introductory notes on why cognitive modeling is important for psychology and cognitive science. Create an `environment.yml` file and add all dependencies we have discussed so far. Then, in addition to the main branch, create separate branches for each of the two team members, from which you will be merging working code into the main branch.

2. Create a *merge conflict* (either for some of the coding exercises or a mock conflict) and resolve it.

3. Explain the differences between the following git commands

   (a.) `git restore`

   (b.) `git checkout`

   (c.) `git reset`

   (d.) `git revert`

   in terms of undoing changes to a repository by providing a minimal (actual or a synthetic) example.

Repository: `https://github.com/Chase-Grajeda/Grajeda-Rowe`
For merge conflict creation and resolution, see commits 2a5852d through b176758

(a.) Git restore provides the option to remove all working changes for staged and unstaged files. Additionally, a user could restore back to a previous version of the file. For restoring staged files, you should use "git restore --staged fname." For restoring a specific version, use "git restore --source commit-id fname."

(b.) Git checkout allows the user to quickly change which branch they are currently observing, whether its any of the branches currently on HEAD or one from the past. Further, it can be used to "restore," or view an old version of a file. If I wanted to quickly switch branches, I would use "git checkout branch-name." If I wanted to view an old version of a file, I would use "git checkout commit-id fname"

(c.) Git reset adjusts history more directly by updating which commit HEAD is currently pointing at. This can be used to completely restore back to a previous commit, but is considered dangerous since history may not be preserved. If I wanted to roll back to commit ABCD123 and history was not a concern, I would use the command "git reset –hard ABCD123."

(d.) Git revert is a safe way to undo changes provided it saves commit history. To accomplish this, a new commit will be created that rewrites your file(s) to a previous version. To do this, use "git revert commit-id"

Ryan Rowe & Chase Grajeda
COGS 4270-01
2024-02-12

Assignment 1 (Non-Programming Questions)

[Q3] Expectations 1 [4 pts]
Suppose that you want to invest some money in the oil market. You believe that the probability of the market going up is 0.8 and the market going down is 0.2. Further, if the market goes up, oil prices will increase by 1%, if it goes down, oil prices will drop by 10%.

What is your expectation? Would you invest in this market? Assuming the increase/drop in prices is fixed, what is the minimal probability of prices going up in order for you to invest rationally (according to expectation) in this market? Discuss a limitation of expectations when making single-shot, real-life decisions.

X_1 = market goes up, X_2 = market goes down.

$$\mathbb{E}[X] = \mathbb{E}[X_1]\mathbb{P}[X_1] + \mathbb{E}[X_2]\mathbb{P}[X_2] = (1\%)(0.8) + (-10\%)(0.2) = 0.8\% + (-2\%) =$$
$$= -1.2\%$$

My expectation is that the market will go down -1.2% on average. Thus, I would not invest in this market, since the smaller chance of loosing a lot outweighs the larger chance of gaining a little.

In order to invest rationally (according to expectation), we would need $\mathbb{E}[X] \geq 0\%$.

Assume only possibilities are X_1 and X_2. Can then set $\mathbb{P}[X_1]+\mathbb{P}[X_2] = 1 \to \mathbb{P}[X_1] = 1-\mathbb{P}[X_2]$

$$\mathbb{E}[X] = \mathbb{E}[X_1]\mathbb{P}[X_1] + \mathbb{E}[X_2]\mathbb{P}[X_2] = \mathbb{E}[X_1]\mathbb{P}[X_1] + \mathbb{E}[X_2](1 - \mathbb{P}[X_1]) =$$
$$= \mathbb{E}[X_1]\mathbb{P}[X_1] - \mathbb{E}[X_2]\mathbb{P}[X_1] + \mathbb{E}[X_2] = (\mathbb{E}[X_1] - \mathbb{E}[X_2])\mathbb{P}[X_1] + \mathbb{E}[X_2] \geq 0\% \longrightarrow$$
$$\longrightarrow \mathbb{P}[X_1] \geq \frac{-\mathbb{E}[X_2]}{\mathbb{E}[X_1] - \mathbb{E}[X_2]} = \frac{-(-10\%)}{(+1\%) - (-10\%)} = \frac{10}{1+10} = \frac{10}{11} = 0.90\overline{90}$$

In order to invest rationally in this market, the probability of the market going up would have to be at least $0.90\overline{90}$.

<u>Limitations of Expectations:</u> One big limitation when it comes to real-life decisions is risk. In this problem, the risk of loosing big if the market goes down may make it not worth it to invest from a personal perspective. Even if the expected value was positive (such as if the probabilities are better than the “rational” ones we calculated), the fact that you could lose a lot, even if you are very likely to gain a little, could change your decision.

Assignment 1 (Non-Programming Questions)

[Q4] Expectations 2 [4 pts]

1. Show the following identity for the variance of a random variable X:

$$\mathrm{Var}[X] = \mathbb{E}\big[X^2\big] - \mathbb{E}[X]^2$$

2. Show the following property for the variance of a random variable X and a scalar α:

$$\mathrm{Var}[\alpha X + \beta] = \alpha^2 \mathrm{Var}[X]$$

3. Assume you are given a random variable X with a standard normal distribution (mean 0, variance 1). We can write this as

$$X \sim \mathrm{Normal}(\mu = 0, \sigma = 1)$$

One way to sample from this distribution is using NumPy's function `numpy.random.randn`. What transformation do you need to apply to the sampled values (i.e., the outputs of the function) such that they are now distributed according to

$$\tilde{X} \sim \mathrm{Normal}(\mu = 3, \sigma = 5)$$

(1.) $\mathrm{Var}[X] =$

$= \mathbb{E}\big[(X - \mathbb{E}[X])^2\big]$

$= \mathbb{E}\big[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2\big]$

$= \mathbb{E}\big[X^2\big] + \mathbb{E}\big[-2X\mathbb{E}[X]\big] + \mathbb{E}\big[\mathbb{E}[X]^2\big]$

$= \mathbb{E}\big[X^2\big] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2\mathbb{E}[1]$

$= \mathbb{E}\big[X^2\big] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2$

$= \mathbb{E}\big[X^2\big] - \mathbb{E}[X]^2$

(2.) $\mathrm{Var}[\alpha X + \beta] =$

$= \mathbb{E}\big[(\alpha X + \beta)^2\big] - \mathbb{E}[\alpha X + \beta]^2$

$= \mathbb{E}\big[\alpha^2 X^2 + 2\alpha\beta X + \beta^2\big] - \big(\alpha\mathbb{E}[X] + \beta\big)^2$

$= \alpha^2\mathbb{E}[X^2] + 2\alpha\beta\mathbb{E}[X] + \beta^2 - \alpha^2\mathbb{E}[X]^2 - 2\alpha\beta\mathbb{E}[X] - \beta^2$

$= \alpha^2\mathbb{E}[X^2] - \alpha^2\mathbb{E}[X]^2$

$= \alpha^2\big(\mathbb{E}\big[X^2\big] - \mathbb{E}[X]^2\big)$

$= \alpha^2\mathrm{Var}[X]$

Assignment 1 (Non-Programming Questions)

(3.) Transform a normal distribution X (with mean $\mu = 0$ and variance $\sigma = 1$) into normal distribution $\tilde{X}$ (with $\tilde{\mu} = 3$ and $\tilde{\sigma} = 5$):

We know (in general) that $\mathbb{E}[X] = \mu$ and $\text{Var}[X] = \sigma^2$. Assuming the transformation takes the form $\tilde{X} = \alpha X + \beta$ (since part 2 used that), we can calculate α and β using the given means and variances.

$$\text{Var}\big[\tilde{X}\big] = \text{Var}\big[\alpha X + \beta\big] = \alpha^2 \text{Var}\big[X\big] \longrightarrow \alpha^2 = \frac{\text{Var}\big[\tilde{X}\big]}{\text{Var}[X]} = \frac{\tilde{\sigma}^2}{\sigma^2} = \left(\frac{\tilde{\sigma}}{\sigma}\right)^2 \longrightarrow \alpha = \frac{\tilde{\sigma}}{\sigma}$$

$$\mathbb{E}\big[\tilde{X}\big] = \mathbb{E}\big[\alpha X + \beta\big] = \alpha \mathbb{E}[X] + \beta \longrightarrow \beta = \mathbb{E}\big[\tilde{X}\big] - \alpha \mathbb{E}[X] = \tilde{\mu} - \alpha\mu = \tilde{\mu} - \frac{\tilde{\sigma}}{\sigma}\mu$$

<u>Generalized transform:</u> $\tilde{X} = \alpha X + \beta = \frac{\tilde{\sigma}}{\sigma} X + \tilde{\mu} - \frac{\tilde{\sigma}}{\sigma}\mu = \frac{\tilde{\sigma}}{\sigma}(X - \mu) + \tilde{\mu} \longrightarrow \tilde{X} = \frac{\tilde{\sigma}}{\sigma}(X - \mu) + \tilde{\mu}$

<u>For this problem specifically:</u> $\frac{5}{1}(X - 0) + 3 = 5X + 3 = \tilde{X}$

Assignment 1 (Non-Programming Questions)

[Q5] Simple Bayesian Inference [4 pts]
The inhabitants of an island tell the truth 1/3 of the time. They lie with probability 2/3. On an occasion, after one of them made a statement, you ask another "Was that statement true?" and he says "yes". What is the probability that the statement was indeed true?

Statement 1	Statement 2 ("yes")	Probability for Both	Is Possible?
True (1/3)	["Yes"] True (1/3)	(1/3)(1/3) = (1/9)	Possible
True (1/3)	["No"] False (2/3)	(1/3)(2/3) = (2/9)	Not
False (2/3)	["No"] True (1/3)	(2/3)(1/3) = (2/9)	Not
False (2/3)	["Yes"] False (2/3)	(2/3)(2/3) = (4/9)	Possible

S_1 can be either True or False. S_2 can be either True or False, but since we know it's always "Yes" it's truth value must be the same as S_1 (as shown by the above table). Thus, we'll instead have S_2 with the possibilities "Yes" or "No".

We have the following probabilities (since $\mathbb{P}[\mathrm{T}] = \frac{1}{3}$ and $\mathbb{P}[\mathrm{F}] = \frac{2}{3}$):
$\mathbb{P}[S_1 = \mathrm{T}] = \frac{1}{3}$, $\mathbb{P}[S_1 = \mathrm{F}] = \frac{2}{3}$, $\mathbb{P}[S_2 = \mathrm{Y} \mid S_1 = \mathrm{T}] = \frac{1}{3}$, $\mathbb{P}[S_2 = \mathrm{Y} \mid S_1 = \mathrm{F}] = \frac{2}{3}$.

We want to calculate the probability that the 1st statement is true: $\mathbb{P}[S_1 = \mathrm{T} \mid S_2 = \mathrm{Y}]$

Bayes' Rule: $\mathbb{P}[S_1 = \mathrm{T} \mid S_2 = \mathrm{Y}] = \dfrac{\mathbb{P}[S_2 = \mathrm{Y} \mid S_1 = \mathrm{T}]\mathbb{P}[S_1 = \mathrm{T}]}{\mathbb{P}[S_2 = \mathrm{Y}]}$

From the above table: $\mathbb{P}[S_2 = Y] = \left(\frac{1}{3} * \frac{1}{3}\right) + \left(\frac{2}{3} * \frac{2}{3}\right) = \frac{1}{9} + \frac{4}{9} = \frac{5}{9}$ (sum of "Yes" rows)

Thus, $\mathbb{P}[S_1 = \mathrm{T} \mid S_2 = \mathrm{Y}] = \dfrac{\mathbb{P}[S_2 = \mathrm{Y} \mid S_1 = \mathrm{T}]\mathbb{P}[S_1 = \mathrm{T}]}{\mathbb{P}[S_2 = \mathrm{Y}]} = \dfrac{(1/3)(1/3)}{(5/9)} = \dfrac{(1/9)}{(5/9)} = \dfrac{1}{5}$

Ryan Rowe & Chase Grajeda
COGS 4270-01
2024-02-12

Assignment 1 (Non-Programming Questions)

[Q6] Murder Mystery Revised [6 pts]
Construct a small story of your own choosing in the spirit of the "Murder Mystery" we considered in class (can be from a completely different domain). Define a "prior" distribution, a "likelihood", and justify your selection of probabilities. Create a first folder in your team repository and write a script containing fully vectorized code which simulates your story and outputs a table representing the approximate joint probabilities (which we approximated through the relative frequencies across N simulation runs). Compare this table to the analytic probabilities. How large should N be for the approximate probabilities to become almost indistinguishable from the analytic ones?

See revised-mystery.ipynb on GitHub for implementation

Assignment 1 (Non-Programming Questions)

[Q7] Priors, Sensitivity, Specificity [6 pts + 4 bonus pts]
Let's revisit the task we disease problem we tackled during class. Imagine you are a medical researcher analyzing the effectiveness of a new diagnostic test for a rare disease **X**. This disease affects 1% of the population. The probability of a true positive (the test correctly identifies an individual as having the disease) is 95%. This is also known as the *sensitivity* of the test. The probability of a true negative (the test correctly identifies an individual as not having the disease) is 90%. This is also known as the *specificity* of the test.

We will now consider a question of sensitivity analysis (not to be confused with the sensitivity of a test): How would the posterior probability change if the prior, the sensitivity, or the specificity of the test were to test. Write a Python program which produces three pretty and annotated 2D graphs depicting

1. The posterior probability (Y-axis) of actually having the disease given a positive test as a function of the prior probability (X-axis), assuming fixed sensitivity and specificity.
2. The posterior probability (Y-axis) of actually having the disease given a positive test as a function of the test's sensitivity (X-axis), assuming fixed prior and specificity.
3. The posterior probability (Y-axis) of actually having the disease given a positive test as a function of the test's specificity (X-axis), assuming fixed prior and sensitivity.

Briefly discuss how the posterior changes as a function of each of the quantities.

Bonus [4 pts] Generate three 3D plots (either surface plots or scatter plots) depicting the same posterior probability as a function of the combination of two quantities (prior - sensitivity, prior - specificity, sensitivity - specificity).

See sens-analysis.ipynb on GitHub for implementation, graphs, and discussion

Assignment 1 (Non-Programming Questions)

[Q8] Monte Carlo Approximation [4 pts]
Write a Python program that approximates the value of π using Monte Carlo approximation. Your program should generate a sequence of random points and use these points to estimate the value of π. The accuracy of the approximation should improve as the number of points increases. Here are some hints:

- Your program should generate random points with x and y coordinates ranging between -1 and 1. This will simulate points within a 2×2 square that circumscribes a unit circle centered at the origin $(0,0)$.
- For each generated point, determine whether it falls inside the unit circle. A point $p = (x, y)$ is inside the circle if $x^2 + y^2 \leq 1$.
- Use the ratio of the number of points that fall inside the circle to the total number of generated points to approximate π.
  The formula is given by $\pi \approx 4 \times \frac{\text{[number of points inside]}}{\text{[total number of points]}}$.

See monte-carlo.ipynb on GitHub for implementation and graphs

Ryan Rowe & Chase Grajeda

Assignment 1 (Non-Programming Questions)

[Q9] AI-Assisted Programming [4 pts]
Use ChatGPT or any other large language model (LLM) to generate a function called `multivariate_normal_density(x, mu, Sigma)` which returns the density of a D-dimensional vector `x` given a D-dimensional mean (location) vector `mu` and a $(D \times D)$-dimensional covariance matrix `Cov`. Compare the outputs of your function with those obtained using SciPy's `scipy.stats.multivariate_normal` for a few parameterizations including a spherical Gaussian (zero covariance, shared variance across dimensions), a diagonal Gaussian (zero covariance, different variance for each dimension), and a full-covariance Gaussian (non-zero covariance, different variance for each dimension). Describe briefly how the LLM performed.

See ai-assisted.ipynb on GitHub for implementation and discussion