1. Consider throwing a 6 sided die two times in succession. Let A be the number of the first throw and B be the number of the second throw. Let S be the random variable that is the sum of the numbers. What is:

a) P(A=3, B=3 | S=6)

b) P(S=6 | A=3, B=3)

c) P(A=1, B=1 | S=6)

d) P(A=1|S=2)

e) P(A=3|S=2)

f) P(A=2|S=6)

g) P (S=12) h) P(S=6)

A, P(A=3, B=3 | S=6)

To find this probability, we need to identify all the combinations where the sum of the numbers is 6 and the first throw is 3 and the second throw is 3. There are indeed four such combinations: (3,3), (1,5), (5,1), (2,4), (4,2).

So, the probability P(A=3, B=3 | S=6) is the probability of getting (3,3) out of these five possible outcomes:

P(A=3, B=3 | S=6) = **1/5.**

B) P(S=6 | A=3, B=3)

A=3 and B=3, the sum is already determined to be 6.

P(S=6 | A=3, B=3) = **1.**

C) P(A=1, B=1 | S=6)

it's not possible for both A and B to be 1 and still have a sum of 6. So this probability is 0:

P(A=1, B=1 | S=6) = **0**

D) P(A=1 | S=2)

To determine the probability that the first throw results in a 1 given that the sum of the numbers is 2.Since the sum is 2, the only possible combination is (1, 1). So, if the sum is 2, it is certain that the first throw is 1.

P(A=1 | S=2) = B

E) P(A=3|S=2)

There is no way to get a sum of 2 with the number 3 on a six-sided die, so the probability is 0.

P(A=3 | S=2) = **0**

F) P(A=2|S=6)

Given that the sum of the numbers is 6, the possibilities are {(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)}. The probability of getting A=2 is 1/5.

P(A=2 | S=6) = **1/5.**

G) P (S=12)

The only way to get a sum of 12 is by rolling both dice as 6, so the probability is 1/36.

P (S=12) = **1/36.**

H) P(S=6)

The possibilities for the sum of 6 are {(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)}. The probability is 1/6 for each pair, so the combined probability is 5/36.

H) P(S=6) = **5/36.**

2, Suppose you roll an N-sided dice, each side labeled with a number 1,2,…,N. Further assume the dice is fair(i.e. any side is equally likely to be rolled). What is the expected number of rolls needed before every side comes up at least once?

Hint: Let the random variable X be the number of rolls needed to see all sides at least once. Let Xi be the number of times you rolled while having seen (i-1) sides. Express X as a function of the Xi's.

Let's define:

X: The total number of rolls needed to see all sides at least once.

Xi: The number of rolls needed to see the ith side for the first time, given that you've already seen (i-1) sides.

We can see that X can be expressed as the sum of the Xi's:

$X = X_1 + X_2 + ... + XN$

Now, what's the expected value of each Xi?

To see the first side (X_1), you only need one roll (since there's no requirement for any other side to appear first). To see the second side (X_2), after seeing the first one, the probability of not seeing the second side on any given roll is (N-1)/N. So, the expected number of additional rolls needed is 1 / (probability of seeing the second side on the next roll) = N/(N-1). This follows a geometric distribution. Similarly, for the third side (X_3), the probability of not seeing it after seeing the first two is (N-2)/N, so the expected number of additional rolls is N/(N-2). We can continue this logic for all remaining sides. Therefore, the expected total number of rolls (X) becomes:

$$X = 1 + N/(N-1) + N/(N-2) + ... + N/1$$

The sum approximates to $\mathbf{X \approx N * \ln(N)}$**.**

3.Consider three binary variables a,b,c ∈ {0,1} having the joint distribution given in the table below. Show by direct computation that this distribution has the property that a and be are marginally dependent, so that p(a,b)≠p(a)p(b), but that they become independent when conditioned on c, so that p(a,b|c)=p(a|c)p(b|c) for both c=0 and c=1

4. Consider these two propositions for conditional independence of X and Y, given Z: (A) P(X|Y,Z)=P(X|Z) or P(Y,Z)=0. (B) P(X,Y|Z)=P(X|Z)P(Y|Z) Show that (A) implies (B) and that (B) implies (A). The point of this problem is to get used to manipulating expressions using basic probability definitions.

To show that proposition (A) implies proposition (B) and vice versa, let's first prove (A) implies (B).

(A) P(X | Y , Z) = P(X | Z) or P(Y, Z) = 0

From proposition (A), we can write:

P(X,Y|Z) = P(X|Y,Z) * P(Y|Z)   -> by definition of conditional probability

Substitute from proposition (A):

P(X,Y|Z) = P(X|Z) * P(Y|Z)   -> since P(X|Y,Z) = P(X | Z)

This is exactly what proposition (B) states, therefore proving that (A) implies (B). Now, to show that proposition (B) implies proposition (A):

(B) P(X,Y|Z) = P(X|Z) * P(Y|Z)

P(X|Y,Z) = P(X,Y|Z) / P(Y|Z) -> Using the definition of conditional probability:

Substitute from proposition (B) into the equation above:

P(X|Y,Z) = [P(X|Z) * P(Y|Z)] / P(Y|Z)

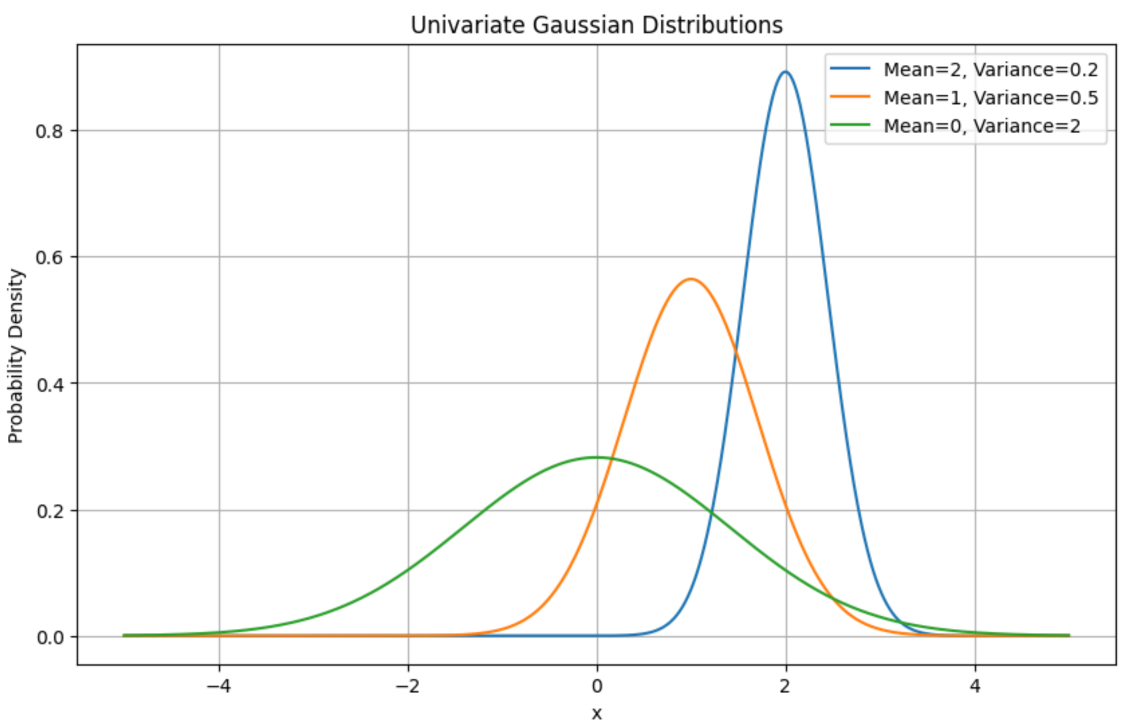
P(X|Y,Z) = P(X|Z)

This proves that (B) implies (A).

Thus, we've shown that proposition (A) implies proposition (B) and that proposition (B) implies proposition (A), demonstrating the relationship between the two conditional independence propositions.

5. For this problem you should code up the formula for the univariate and bivariate Gaussian distributions rather using existing implementation (python). You are welcome to use such functions to check your answer. The reason for entering the formula yourself is to get some feel for the form of the function. For the multivariate version, you need to find the 2 | Page determinant of a matrix, which is non-trivial in general, but very easy for 2x2, where it is given by det(A)=A11A22-A12A21. You can also make use of an external function for the inverse of a matrix, but again, for 2x2 it is a simple formula that you can look up.

a) Plot three univariate Gaussians on one graph. Specifically, plot Gaussians with means and variances (2,0.2), (1, 0.5), and (0, 2). Be sure to provide a caption.

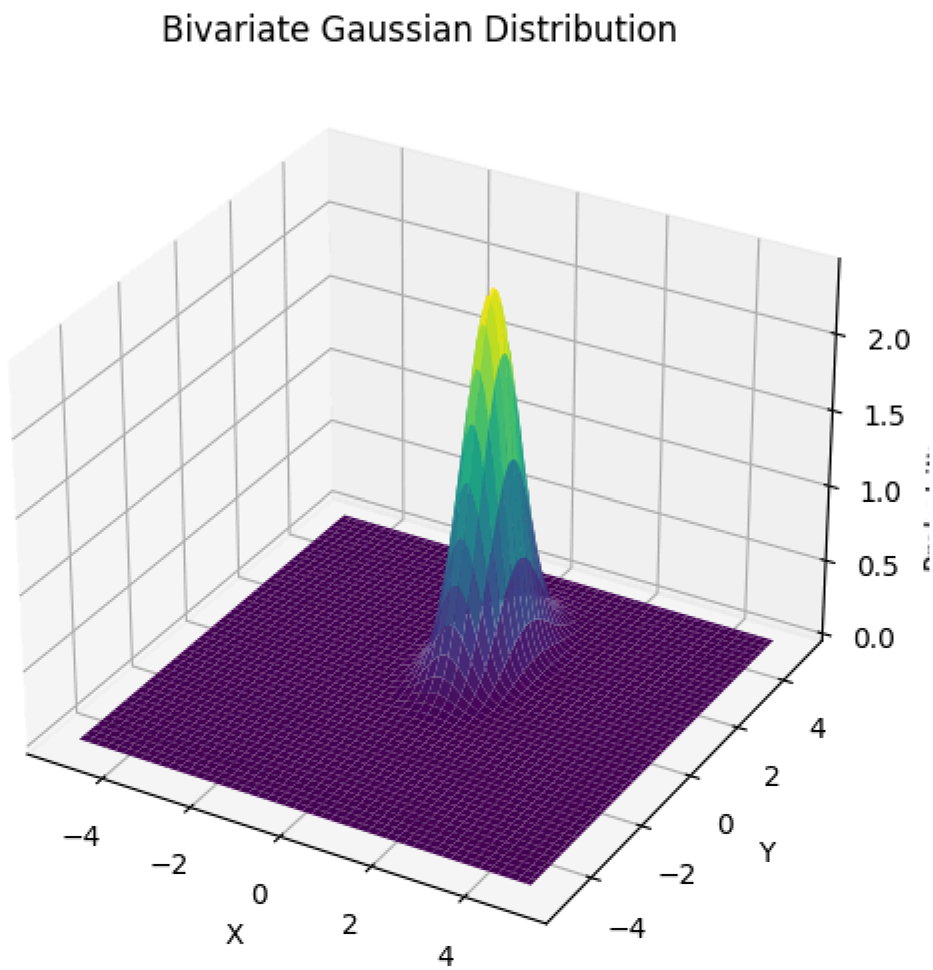


b) Make a 3D plot of p(x,y) for a bivariate Gaussian with mean (0,0) and covariance matrix. 0.5 0.8 0.8 2.0

[[0.5 0.8]

[0.8 2.0]].

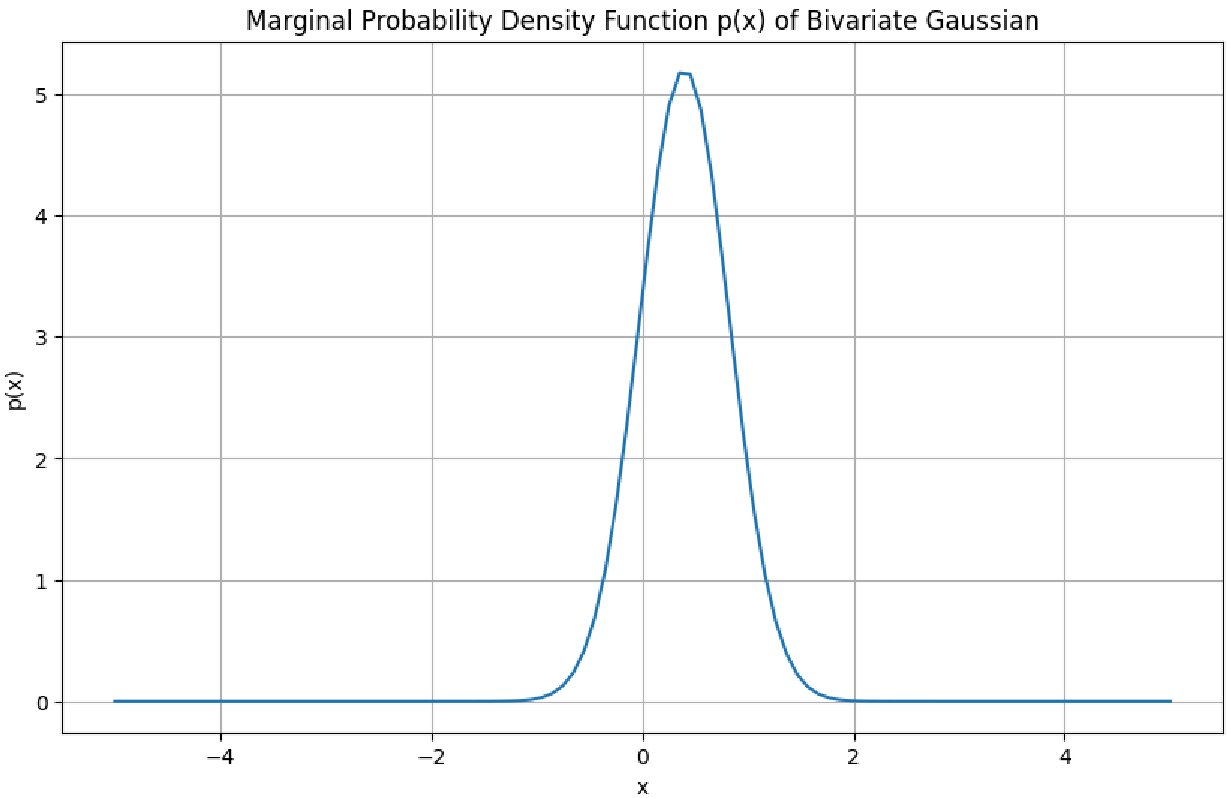
Be sure to provide a caption.



For problems 6 and 7 , the main point is getting familiar with the joint distribution for continuous domains, andalso seeing two very interesting properties about Gaussian distributions in action. Once you have done the first one,the second one should not take too much time as you will have built some of the programming infrastructure.

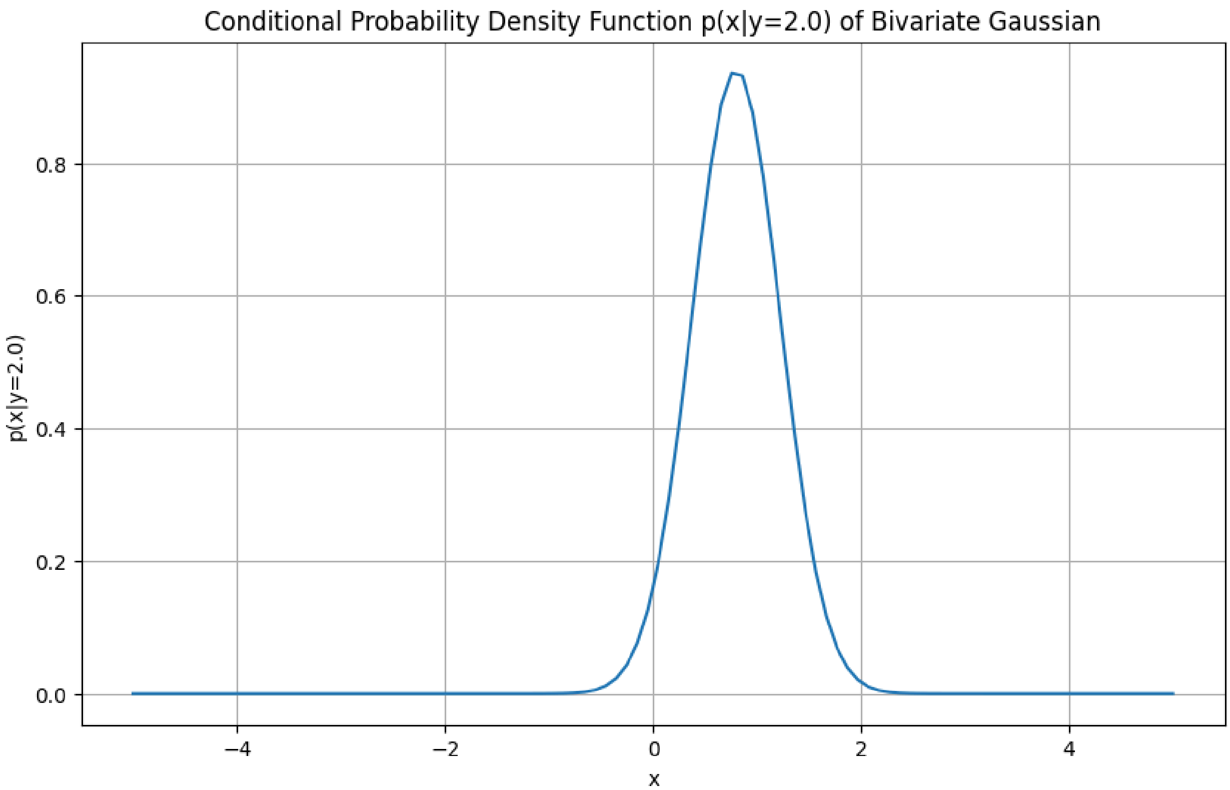
6. For the bivariate Gaussian in part (b) of problem 5, using numerical integration, approximate p(x) for a sensible stepping of x values and plot the result. Be sure to provide a caption. Does it have the shape you expect, and what is that shape?.

Alternative: If proof by programming is not for you, you can instead show analytically that for a bivariate Gaussian p(x,y), p(x) is also Gaussian. What do you expect the area of the curve to be? It may be easier to work with the precision matrix which is the inverse of the covariance matrix and which is also symmetric.



7. For the bivariate Gaussian in part (b) of problem 5, plot p(x|y=2.0) . Be sure to provide a good caption.Does it have the shape you expect, and what is that shape? What do you expect the area of the curve to be?

Alternative: If proof by programming is not for you, show analytically that for a bivariate Gaussian p(x,y),p(x|y) is also Gaussian . It may be easier to work with the precision matrix which is the inverse of the covariance matrix and which is also symmetric.


Conditional Probability Density Function p(x|y=2.0) of Bivariate Gaussian
p(x|y=2.0)
0.0
0.2
0.4
0.6
0.8
−4
−2
0
2
4
x