

Homework 1 - Probability Review

Darren Biskup

2025-09-10

1. Use function `random()` from the `random` Python module to estimate the value of π . Your algorithm must be based on the given random generator (you may not use any alternative algorithm for generating the value of π) and you may not use any other package/library.

First we can spawn 1000 trials of uniformly random points in a square of side length 1 centered at the origin. Then if we count how many of those points fall within a circle of radius 0.5, we can use the ratio of points in the circle to total points to estimate π . (law of large numbers)

The theoretical probability that a randomly sampled point in the unit square is also inside the circle is the ratio of the area of the circle to the area of the square.

$$A_{circle} = \pi r^2 = \pi(0.5)^2 = \frac{\pi}{4}$$
$$A_{square} = 1^2 = 1$$

Therefore the probability that the point is in the circle is: $\frac{\pi}{4}$. The law of large numbers means that if we conduct an experiment where we sample a point in the square many times, the fraction of points that fall in the circle will converge to $\frac{\pi}{4}$. We can then multiply our estimate of this probability by 4 to get an estimate of π .

First spawn 1000 random points in the square:

```
import numpy as np
import matplotlib.pyplot as plt

num_trials = 100000
x = np.random.rand(num_trials) - 0.5
y = np.random.rand(num_trials) - 0.5
```

Next we can count how many of those points fall within the circle and divide by the total number of points.

```
num_inside_circle = np.sum(x**2 + y**2 <= 0.25)
pi_estimate = (num_inside_circle / num_trials) * 4
print(f"Estimated value of pi: {pi_estimate}")
```

Estimated value of : 3.14504

2. Covariance matrix for a multivariate random variable is defined as $\sum_{xx} = E[(x-\mu)(x-\mu)^T]$. The cross covariance matrix for two multivariate random variables is defined as $\sum_{xy} = E[(x-\mu_x)(y-\mu_y)^T]$.

a. Express the covariance of a sum of two multivariate random variables in terms of covariance and cross-covariance.

Let a new random variable Z be the sum of 2 multivariate random variables X and Y . Then the covariance of Z is:

$$\Sigma_{zz} = E[(Z - \mu_z)(Z - \mu_z)^T]$$

If we expand the outer product and apply properties of expectation we get:

$$\begin{aligned}\Sigma_{zz} &= E[ZZ^T - Z\mu_z^T - \mu_z Z^T + \mu_z \mu_z^T] \\ &= E[ZZ^T] - E[Z]\mu_z^T - \mu_z E[Z]^T + \mu_z \mu_z^T \\ &= E[ZZ^T] - \mu_z \mu_z^T - \mu_z \mu_z^T + \mu_z \mu_z^T \\ &= E[ZZ^T] - \mu_z \mu_z^T\end{aligned}$$

now we can substitute $Z = X + Y$ and $\mu_z = \mu_x + \mu_y$:

$$\begin{aligned}\Sigma_{zz} &= E[(X + Y)(X + Y)^T] - (\mu_x + \mu_y)(\mu_x + \mu_y)^T \\ &= E[XX^T + XY^T + YX^T + YY^T] - (\mu_x \mu_x^T + \mu_x \mu_y^T + \mu_y \mu_x^T + \mu_y \mu_y^T) \\ &= E[XX^T] + E[XY^T] + E[YX^T] + E[YY^T] - \mu_x \mu_x^T - \mu_x \mu_y^T - \mu_y \mu_x^T - \mu_y \mu_y^T \\ &= (E[XX^T] - \mu_x \mu_x^T) + (E[XY^T] - \mu_x \mu_y^T) + (E[YX^T] - \mu_y \mu_x^T) + (E[YY^T] - \mu_y \mu_y^T)\end{aligned}$$

therefore:

$$\Sigma_{zz} = \Sigma_{xx} + \Sigma_{xy} + \Sigma_{yx} + \Sigma_{yy}$$

b. Prove that the cross-covariance of two independent multivariate random variables is a zero-matrix.

the cross covariance of two multivariate random variables as stated is:

$$\begin{aligned}\Sigma_{xy} &= E[(X - \mu_x)(Y - \mu_y)^T] \\ &= E[XY^T] - \mu_x \mu_y^T\end{aligned}$$

if X and Y are independent, then $E[XY^T] = E[X]E[Y]^T = \mu_x \mu_y^T$. This is because $p(x, y) = p(x)p(y)$:

$$\begin{aligned}E[XY^T] &= \int_x \int_y xy^T p(x, y) dy dx \\ &= \int_x \int_y xy^T p(x) p(y) dy dx \\ &= \int_x xp(x) \left(\int_y y^T p(y) dy \right) dx \\ &= \mu_y^T \int_x xp(x) dx \\ &= \mu_x \mu_y^T\end{aligned}$$

Therefore:

$$\Sigma_{xy} = \mu_x \mu_y^T - \mu_x \mu_y^T = 0$$

c. What is the corollary of the above two results?

When X and Y are independent, the covariance of their sum is thus:

$$\begin{aligned}\Sigma_{zz} &= \Sigma_{xx} + \Sigma_{xy} + \Sigma_{yx} + \Sigma_{yy} \\ \Sigma_{zz} &= \Sigma_{xx} + 0 + 0 + \Sigma_{yy} \\ \Sigma_{zz} &= \Sigma_{xx} + \Sigma_{yy}\end{aligned}$$

3. A robot has 10 (unreliable) sensors that independently detect the presence of an object in its intake mechanism. The probability of detecting the object for each sensor is 0.1.

a. Write the distribution of the random variable that represents how many sensors have detected the presence of an object?

Let X be the random variable representing how many sensors currently detect an object. If each sensor has an independent probability of 0.1 of detecting an object, then X follows a binomial distribution with $n = 10$ and $p = 0.1$:

$$X \sim \text{Bin}(n = 10, p = 0.1)$$

b. What is the probability that at least one sensor has detected the presence of an object?

The probability that at least one sensor has detected an object is equal to one minus the probability that no sensors have detected an object. If each sensor has a 90 percent chance of not detecting an object, then the probability that all 10 sensors do not detect an object is:

$$P(X = 0) = 0.9^{10}$$

so the probability that at least one sensor detects an object is:

$$1 - P(X = 0) = 1 - 0.9^{10} \approx 0.65132$$

4. A two-dimensional normally-distributed random variable has the uncertainty ellipse centered at (3, 4) and one principal axis of the ellipse forms a 30-degree angle with the x-axis of the coordinate system. The variance associated with that axis is 1. The variance associated with the other axis is 0.25. Write the probability density function of this random variable.

We are told that the random variable is normally distributed with mean $\mu = (3, 4)$. This means that the PDF will be in a gaussian form:

$$f(x) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$$

To complete the PDF we need to find the covariance matrix Σ and the mean vector μ . The problem statement gives us the mean vector directly as $\mu = [3 \ 4]^T$.

The covariance matrix for a 2D gaussian in the original coordinates will take on the form:

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \rho_{12}\sigma_{11}\sigma_{22} \\ \rho_{12}\sigma_{11}\sigma_{22} & \sigma_{22}^2 \end{bmatrix}$$

But the problem gives us the variances along the principal axes of the ellipse. So it is more convenient to express the covariance matrix as a rotation of the covariance in its principal axis frame:

$$\Sigma = RDR^T$$

Where D is the diagonal matrix of variances along the principal axes, and R is the rotation matrix that aligns the principal axes with the coordinate axes. We are told that the variances along the principal axes are 1 and 0.25 so $D = \begin{bmatrix} 1 & 0 \\ 0 & 0.25 \end{bmatrix}$. We are also told that the principal

axis associated with variance 1 forms a 30 degree angle with the original coordinates' x-axis, so the rotation matrix is $R = \begin{bmatrix} \cos(30^\circ) & -\sin(30^\circ) \\ \sin(30^\circ) & \cos(30^\circ) \end{bmatrix}$.

Thus:

$$\Sigma = RDR^T = \begin{bmatrix} \cos(30^\circ) & -\sin(30^\circ) \\ \sin(30^\circ) & \cos(30^\circ) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.25 \end{bmatrix} \begin{bmatrix} \cos(30^\circ) & \sin(30^\circ) \\ -\sin(30^\circ) & \cos(30^\circ) \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 0.8125 & 0.1875\sqrt{3} \\ 0.1875\sqrt{3} & 0.4375 \end{bmatrix}$$

Now I can use `sympy` to compute the determinant and inverse of Σ , and write out the full PDF:

```
from sympy import *

theta = rad(30)
mu = Matrix([3,4])
R = Matrix([[cos(theta), -sin(theta)], [sin(theta), cos(theta)]])
D = Matrix([[1,0], [0,0.25]])

Sigma = R * D * R.T
Sigma_det = Sigma.det()
Sigma_inv = Sigma.inv()

x1, x2 = symbols('x_1 x_2')

pdf = (
    1/(2*pi*sqrt(Sigma_det)) *
    exp((-1/2 * (Matrix([x1, x2]) - mu).T * Sigma_inv * (Matrix([x1, x2]) - mu))[0])
)

pdf
```

$$\frac{1.0e^{(x_1-3)(-0.875x_1-0.75\sqrt{3}(2.0-0.5x_2)+2.625)+(x_2-4)(-1.625x_2-0.75\sqrt{3}(1.5-0.5x_1)+6.5)}}{\pi}$$

Final PDF:

$$f(x_1, x_2) = \frac{1}{\pi} \exp \left[(x_1 - 3) \left(-0.875x_1 - 0.75\sqrt{3}(2.0 - 0.5x_2) + 2.625 \right) \right. \\ \left. + (x_2 - 4) \left(-1.625x_2 - 0.75\sqrt{3}(1.5 - 0.5x_1) + 6.5 \right) \right]$$

5. Suppose there is a village in which every family wants to have a boy. If they get a boy, they stop having more children, if they get a girl, they continue to have more children. What is the average number of children per family?

Let X be the random variable representing the number of children in a family. The number of children a family has can be modeled as following a geometric distribution with $p = 0.5$. The average number of children per family is then just the expected value of X :

$$X \sim \text{Geom}(p = 0.5) E[X] = \frac{1}{p} = \frac{1}{0.5} = 2$$

We can also verify this result by considering the possible outcomes. Half of the families will have 1 child (boy on first try). Then a quarter will have 2 children (boy on second try). An eighth will get 3 children (boy on 3rd try), and so on. So,

$$E[X] = \sum_{k=1}^{\infty} k \left(\frac{1}{2}\right)^k$$

this is a known power series which converges to 2. Thus the average number of children per family is 2.