Exercise Walkthrough:
Analysis of a Uniform Distribution on a Disc

Justin Lanfermann

25. June 2025

Overview

This document provides a detailed, step-by-step walkthrough for an exercise on a two-dimensional random variable. The goal is to carefully dissect each part of the problem, explaining not just *what* we are doing, but *why* we are doing it, grounding our reasoning in the definitions and theorems from the "Discrete Probability Theory" script by Niki Kilbertus.

The exercise focuses on a random variable (X, Y) that is uniformly distributed on a unit disc. This scenario is a perfect illustration of how to work with continuous joint distributions and highlights the important distinction between uncorrelated and independent random variables.

Exercise. *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space, and* (X, Y) *a* $\mathbb{R}^2$*-valued RV that is uniformly distributed on the disc* $D_2 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$*. The joint probability density function (pdf) of this distribution is given by:*

$$p_{X,Y}(x, y) = \frac{1}{\pi}\chi_{D_2}(x, y)$$

where χ_{D_2} *is the indicator function for the disc.*

(i) Compute the marginal densities $p_X(x)$ *and* $p_Y(y)$*.*

(ii) Compute the means and variances $\mathbb{E}[X], \mathbb{E}[Y], var[X], var[Y]$*.*

(iii) Check whether $X \perp Y$ *(independent) and whether* X *and* Y *are uncorrelated.*

1 Step-by-Step Solution

1.1 Part (i): Computing the Marginal Densities

Goal: Our first task is to find the individual probability density functions for X and Y, which are called marginal densities. Conceptually, you can think of this as "squashing" the 2D probability mass of the disc onto the x-axis to find the density of X, and onto the y-axis for the density of Y.

Method: We use the formula for marginalization from the script (**Theorem 1.63 (iii)** [1]). For a continuous random variable (X, Y) with joint pdf $p_{X,Y}(x, y)$, the marginal pdf of X is found by integrating over all possible values of Y:

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y)\, dy$$

Execution for $p_X(x)$: The joint pdf $p_{X,Y}(x, y)$ is $\frac{1}{\pi}$ inside the unit disc D_2 and 0 everywhere else. The disc is defined by $x^2 + y^2 \leq 1$. For a fixed value of x, the variable y is constrained by $y^2 \leq 1 - x^2$, which means $-\sqrt{1 - x^2} \leq y \leq \sqrt{1 - x^2}$. This also implies that x must be in the interval $[-1, 1]$, otherwise $\sqrt{1 - x^2}$ is not a real number. For any x outside $[-1, 1]$, the density $p_X(x)$ is 0.

For $x \in [-1, 1]$, we compute the integral:

$$\begin{aligned} p_X(x) &= \int_{-\infty}^{\infty} \frac{1}{\pi} \chi_{D_2}(x, y)\, dy \\ &= \frac{1}{\pi} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1\, dy \\ &= \frac{1}{\pi} [y]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \\ &= \frac{1}{\pi} \left(\sqrt{1-x^2} - (-\sqrt{1-x^2}) \right) \\ &= \frac{2}{\pi} \sqrt{1-x^2} \end{aligned}$$

Combining this with the condition on x, the full marginal density is:

$$p_X(x) = \frac{2}{\pi} \sqrt{1-x^2} \chi_{[-1,1]}(x)$$

Execution for $p_Y(y)$: We could repeat the same calculation for $p_Y(y)$. However, notice that the problem is perfectly symmetric with respect to X and Y. The definition of the disc $x^2 + y^2 \leq 1$ remains the same if we swap x and y. Therefore, the marginal density for Y must have the same functional form as for X.

$$p_Y(y) = \frac{2}{\pi} \sqrt{1-y^2} \chi_{[-1,1]}(y)$$

Quick Check: Does our result for $p_X(x)$ make sense? The function $\frac{2}{\pi}\sqrt{1-x^2}$ looks like a semi-ellipse. It has its maximum value at $x = 0$ and is zero at $x = \pm 1$. This is intuitive: if you slice the disc vertically, the longest slice (and thus most probability mass) is at the center ($x = 0$), and the slices become infinitesimally small as you approach the edges ($x = \pm 1$).

1.2 Part (ii): Computing Means and Variances

Goal: Now we compute the expected value (mean) and variance for both X and Y. These are fundamental properties describing the center and spread of the distributions.

Method for Mean $\mathbb{E}[X]$: According to **Definition 2.1**, the expectation of a continuous RV is:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot p_X(x)\, dx$$

Execution for $\mathbb{E}[X]$:

$$\mathbb{E}[X] = \int_{-1}^{1} x \cdot \left(\frac{2}{\pi} \sqrt{1-x^2} \right) dx = \frac{2}{\pi} \int_{-1}^{1} x\sqrt{1-x^2}\, dx$$

We can solve this integral directly, but it's faster to use a symmetry argument. The function $f(x) = x\sqrt{1-x^2}$ is an odd function ([2]), because x is odd and $\sqrt{1-x^2}$ is even. The integral of any odd function over a symmetric interval like $[-1, 1]$ is zero.

$$\mathbb{E}[X] = 0$$

By symmetry, we immediately know that $\mathbb{E}[Y] = 0$.

Method for Variance var$[X]$: From **Remark 2.6**, the most convenient formula for variance is $\text{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. Since we found $\mathbb{E}[X] = 0$, this simplifies to $\text{var}[X] = \mathbb{E}[X^2]$. To find $\mathbb{E}[X^2]$, we use the Law of the Unconscious Statistician (LOTUS), as stated in **Lemma 2.2** ([3]):

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 \cdot p_X(x)\, dx$$

Execution for $\mathbb{E}[X^2]$:

$$\mathbb{E}[X^2] = \int_{-1}^{1} x^2 \cdot \left(\frac{2}{\pi}\sqrt{1-x^2}\right) dx = \frac{2}{\pi}\int_{-1}^{1} x^2\sqrt{1-x^2}\, dx$$

This integral is not trivial. A standard technique for expressions involving $\sqrt{a^2 - x^2}$ is a trigonometric substitution ([4]). Let $x = \sin(u)$, then $dx = \cos(u)\, du$. The limits of integration change from $x \in [-1, 1]$ to $u \in [-\pi/2, \pi/2]$.

$$\begin{aligned}
\mathbb{E}[X^2] &= \frac{2}{\pi}\int_{-\pi/2}^{\pi/2} \sin^2(u)\sqrt{1-\sin^2(u)} \cdot \cos(u)\, du \\
&= \frac{2}{\pi}\int_{-\pi/2}^{\pi/2} \sin^2(u)\sqrt{\cos^2(u)} \cdot \cos(u)\, du \\
&= \frac{2}{\pi}\int_{-\pi/2}^{\pi/2} \sin^2(u)\cos^2(u)\, du \quad (\text{since } \cos(u) \geq 0 \text{ on } [-\pi/2, \pi/2])
\end{aligned}$$

Now, we use the trigonometric identities $\sin(u)\cos(u) = \frac{1}{2}\sin(2u)$ and $\sin^2(\theta) = \frac{1}{2}(1-\cos(2\theta))$:

$$\begin{aligned}
\mathbb{E}[X^2] &= \frac{2}{\pi}\int_{-\pi/2}^{\pi/2} \left(\frac{1}{2}\sin(2u)\right)^2 du = \frac{2}{\pi}\cdot\frac{1}{4}\int_{-\pi/2}^{\pi/2} \sin^2(2u)\, du \\
&= \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} \frac{1}{2}(1-\cos(4u))\, du \\
&= \frac{1}{4\pi}\left[u - \frac{1}{4}\sin(4u)\right]_{-\pi/2}^{\pi/2} \\
&= \frac{1}{4\pi}\left(\left(\frac{\pi}{2} - 0\right) - \left(-\frac{\pi}{2} - 0\right)\right) = \frac{1}{4\pi}(\pi) = \frac{1}{4}
\end{aligned}$$

So, $\text{var}[X] = \mathbb{E}[X^2] = \frac{1}{4}$. By symmetry, $\text{var}[Y] = \frac{1}{4}$.

1.3 Part (iii): Independence and Correlation

Goal: We need to determine if X and Y are independent and if they are uncorrelated. This is the most insightful part of the exercise.

Method for Independence: According to **Definition 1.72 (iii)**, two continuous random variables X and Y are independent if and only if their joint pdf factorizes into the product of their marginal pdfs for (almost) all (x, y):

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$$

Execution for Independence: Let's compute the product of the marginals we found:

$$p_X(x) \cdot p_Y(y) = \left(\frac{2}{\pi}\sqrt{1-x^2}\right) \cdot \left(\frac{2}{\pi}\sqrt{1-y^2}\right) = \frac{4}{\pi^2}\sqrt{(1-x^2)(1-y^2)}$$

This product is non-zero for any (x, y) in the square $(-1, 1) \times (-1, 1)$. The original joint density is $p_{X,Y}(x, y) = \frac{1}{\pi}$ inside the disc and 0 outside. Clearly, $p_{X,Y}(x, y) \neq p_X(x)p_Y(y)$.

To make this concrete, let's pick a point. Consider $(x, y) = (0.7, 0.7)$. This point is inside the square $(-1, 1) \times (-1, 1)$, so $p_X(0.7)p_Y(0.7) > 0$. However, $x^2 + y^2 = 0.49 + 0.49 = 0.98 \leq 1$, so this point is inside the disc D_2. Here $p_{X,Y}(0.7, 0.7) = 1/\pi$. The values are not equal. A stronger argument: consider the point $(x, y) = (0.8, 0.8)$. Here $x^2 + y^2 = 0.64 + 0.64 = 1.28 > 1$, so the point is outside the disc.

- $p_{X,Y}(0.8, 0.8) = 0$ (since it's outside the disc).
- $p_X(0.8)p_Y(0.8) = \frac{4}{\pi^2}\sqrt{(1 - 0.64)(1 - 0.64)} > 0$.

Since the equality does not hold, X **and** Y **are not independent**. This makes sense: if I tell you $X = 0.9$, you know that Y must be in a very small range around 0. Information about X constrains the possible values of Y.

Method for Correlation: Two variables are uncorrelated if their covariance is zero (**Definition 2.13**). We use the computational formula from **Remark 2.10**:

$$\text{cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Since we already know $\mathbb{E}[X] = 0$ and $\mathbb{E}[Y] = 0$, we just need to compute $\mathbb{E}[XY]$.

$$\mathbb{E}[XY] = \iint_{\mathbb{R}^2} xy \cdot p_{X,Y}(x, y)\, dx\, dy = \iint_{D_2} xy \cdot \frac{1}{\pi}\, dx\, dy$$

Execution for Correlation: This integral is best solved using a change of variables to polar coordinates ([6]), as hinted in the problem. Let $x = r\cos(\theta)$ and $y = r\sin(\theta)$. The Jacobian determinant for this transformation is r. The domain D_2 becomes $r \in [0, 1]$ and $\theta \in [0, 2\pi)$.

$$\begin{aligned}
\mathbb{E}[XY] &= \frac{1}{\pi}\int_0^{2\pi}\int_0^1 (r\cos\theta)(r\sin\theta) \cdot r\, dr\, d\theta \\
&= \frac{1}{\pi}\int_0^{2\pi}\int_0^1 r^3 \cos\theta\sin\theta\, dr\, d\theta \\
&= \frac{1}{\pi}\left(\int_0^1 r^3\, dr\right)\left(\int_0^{2\pi} \cos\theta\sin\theta\, d\theta\right)
\end{aligned}$$

Let's focus on the θ integral. Using the identity $\sin(2\theta) = 2\sin\theta\cos\theta$:

$$\int_0^{2\pi} \cos\theta\sin\theta\, d\theta = \frac{1}{2}\int_0^{2\pi} \sin(2\theta)\, d\theta = \frac{1}{2}\left[-\frac{1}{2}\cos(2\theta)\right]_0^{2\pi} = -\frac{1}{4}(\cos(4\pi) - \cos(0)) = -\frac{1}{4}(1-1) = 0$$

Since the θ integral is 0, the entire expression becomes 0.

$$\mathbb{E}[XY] = 0$$

Therefore, $\text{cov}[X, Y] = 0 - (0)(0) = 0$. This means X **and** Y **are uncorrelated**.

2 Summary and Key Takeaways

We have successfully analyzed the random variable (X, Y) on the unit disc.

- **Marginal Densities:** $p_X(x) = \frac{2}{\pi}\sqrt{1 - x^2}\chi_{[-1,1]}(x)$ and $p_Y(y)$ is analogous.
- **Mean and Variance:** $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ and $\text{var}[X] = \text{var}[Y] = \frac{1}{4}$.

- **Dependence:** X and Y are **uncorrelated** but **not independent**.

The most important lesson from this exercise is the demonstration that **uncorrelated does not imply independent** ([5]).

- **Correlation** measures the strength and direction of a *linear* relationship between two variables. Since the covariance is zero, there is no linear association between X and Y.
- **Independence** is a much stronger condition. It means that knowledge of one variable provides no information whatsoever about the other. In our case, the circular boundary of the support domain D_2 creates a non-linear relationship. Knowing X restricts the possible values of Y, so they are dependent.

This example is a fundamental counterexample to keep in mind throughout your study of probability and statistics.

Explanatory Notes

[1] Marginal Density (Theorem 1.63 (iii)): The marginal pdf of one variable in a multivariate distribution represents the probability distribution of that variable alone. It is obtained by "integrating out" or "summing out" all other variables. For a 2D continuous distribution $p_{X,Y}(x, y)$, this means $p_X(x) = \int p_{X,Y}(x, y)\, dy$.

[2] Odd and Even Functions: A function f is **even** if $f(-x) = f(x)$ for all x. Its graph is symmetric about the y-axis. A function f is **odd** if $f(-x) = -f(x)$ for all x. Its graph has rotational symmetry about the origin. A key property is that for any odd function f, the integral over a symmetric interval is zero: $\int_{-a}^{a} f(x)\, dx = 0$. In our case, the integrand for $\mathbb{E}[X]$ was $x\sqrt{1-x^2}$. Since x is odd and $\sqrt{1-x^2}$ is even, their product is odd, making the integral zero.

[3] Variance and LOTUS (Remark 2.6, Lemma 2.2): The variance, $\mathrm{var}[X]$, measures the spread of a distribution. The formula $\mathrm{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ is often easier for computation than the definition $\mathbb{E}[(X - \mathbb{E}[X])^2]$. To compute $\mathbb{E}[g(X)]$ for some function g, the Law of the Unconscious Statistician (LOTUS) allows us to compute $\int g(x)p_X(x)\, dx$ without first having to find the pdf of the new random variable $Z = g(X)$.

[4] Trigonometric Substitutions and Identities: When an integral contains a term of the form $\sqrt{a^2 - x^2}$, the substitution $x = a\sin(u)$ is often very effective. It simplifies the square root using the identity $1 - \sin^2(u) = \cos^2(u)$. The identities used in the calculation were:

- $\sin^2(u)\cos^2(u) = (\sin(u)\cos(u))^2 = (\frac{1}{2}\sin(2u))^2 = \frac{1}{4}\sin^2(2u)$
- $\sin^2(\theta) = \frac{1}{2}(1 - \cos(2\theta))$

[5] Independence vs. Uncorrelation (Definition 1.72, 2.13): This is a critical distinction.

- **Independence**: $p_{X,Y}(x, y) = p_X(x)p_Y(y)$. This means the distributions are completely unrelated.
- **Uncorrelation**: $\mathrm{cov}[X, Y] = 0$. This only means there is no *linear* trend between the variables.

Independence is a stronger condition. It is always true that **independence implies uncorrelation**. However, as this exercise shows, the reverse is not true. A non-linear relationship can exist between variables that are uncorrelated.

[6] Change of Variables to Polar Coordinates (Prop. 2.51): When dealing with circular domains, switching from Cartesian coordinates (x, y) to polar coordinates (r, θ) simplifies the integration boundaries. The transformation is:

$$x = r\cos(\theta), \quad y = r\sin(\theta)$$

When performing this change in a double integral, we must replace the area element $dx\, dy$ with $|\det(J)|\, dr\, d\theta$, where J is the Jacobian matrix of the transformation. For polar coordinates, this determinant is famously r, so we replace $dx\, dy$ with $r\, dr\, d\theta$.