Project 1B Draft 2: Two-Dimensional Distributions, Marginals and Covariance Structure

Liu Xiaolong   ID: 201900170025

April 7, 2022

Abstract

In this project, I will discuss some special properties about the two-dimensional distributions, such as how to construct a bivariate distribution from two single distributions, what difference between dimension two and higher dimension and some special bivariant distributions. For the basic, we refer [5], [11], [9], [1], [10].

Contents

1 Introduction

In this project, I will discuss some special properties about the two-dimensional distributions, such as how to construct a bivariate distribution from two single disttibutions and what difference between dim $= 2$ and higher dimension and some special bivariate distributions.

For the first question, I will introduce the FGM copula and derive some properties about it. A copula is a function that makes marginals F_X and F_Y to some joint distribution F. It was first introduced by Sklar in [6]. Now copulas and several parametric families of copulas have been widely used in statistics. One of the most popular parametric families is the Farlie-Gumbel-Morgenstern (FGM) copula, whose properties were discussed by Farlie [4].

For the second, we will learn something about the random walk and prove that the simple symmetric random walk on $\mathbb{Z}^d$ is recurrent in dimensions $d \leq 2$ and transient in dimensions $d \geq 3$. The motivation comes from observations of various random motions in physical and biological sciences such as the famous brownian motion.

For the third, we will explore the bivariate normal distribution, the most famous and the most important multivariable distribution in probability theory.

2 Preliminaries: Random Variables, Distributions and Covariance

For now I will introduce some basic definitions which we will use.

Definition 2.1. *Consider random vector $\xi(\omega) = (\xi_1(\omega), \cdots, \xi_n(\omega))$ whose components are random variables in some probability space $(\Omega, \mathscr{A}, P)$. The function $F(x_1, \cdots, x_n) = \mathsf{P}(\xi_1(\omega) < x_1, \cdots, \xi_n(\omega) < x_n)$ is called **distribution function** of random vector ξ.*

In the continuous case, there exists a non-negative function $p(x_1, \cdots, x_n)$ such that

$$F(x_1, \cdots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p(y_1, \cdots, y_n) dy_1 \cdots dy_n,$$

*where $\int_{\mathbb{R}^n} p(y_1, \cdots, y_n) dy_1 \cdots dy_n = 1$ and it is called **probability density function**.*

Definition 2.2. *The random variables $\xi_1, \cdots, \xi_n$ are said to be **(mutually) independent** if*

$$\mathsf{P}(\xi_1(\omega) = x_1, \cdots, \xi_n(\omega) = x_n) = \mathsf{P}(\xi_1(\omega) = x_1) \cdots \mathsf{P}(\xi_n(\omega) = x_n))$$

for all $x_1, \cdots, x_n$.

Example 1 (Multinomial distribution)**.** *We let the possibly results of the test are $A_1, ..., A_r$ where $P(A_i) = p_i$ and $p_1 + ... + p_r = 1$. We repeat it n times and assume the results are independent. If we let ξ_i be the number of occurrences of A_i, then*

$$\mathsf{P}(\xi_i = k_i) = \frac{n!}{k_1! \cdots k_r!} p_1^{k_1} \cdots p_r^{k_r};$$

From now we just discuss marginal distribution in the case of two-dimensional distributions and the higher dimension are similar.

Definition 2.3 (Marginal)**.** *We consider two-dimensional distributions (ξ, η).*

(1)[General] Let its distribution function is $F(x,y)$ and let $F_1(x) = \mathsf{P}(\xi < x) = F(x,+\infty)$ and $F_2(y) = \mathsf{P}(\eta < y) = F(+\infty, y)$. These are called the ***marginal distribution functions*** *of $F(x,y)$.*

(2)[Discrete] If ξ take values at $x_1, x_2, ...$ and η take values at $y_1, y_2, ...$, and if we let $\mathsf{P}(\xi = x_i, \eta = y_j) = p(x_i, y_j)$ and $\mathsf{P}(\xi = x_i) = p_1(x_i), \mathsf{P}(\eta = y_j) = p_2(y_j)$, then

$$\sum_j p(x_i, y_j) = p_1(x_i), \sum_i p(x_i, y_j) = p_2(y_j),$$

and these are marginal distribution functions;

(3)[Continuous] Let its probability density function is $p(x,y)$,then

$$F_1(x) = \int_{-\infty}^{x} du \int_{-\infty}^{\infty} p(u,y)dy, F_2(y) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{y} p(x,v)dv,$$

so the probability density functions of $F_1(x), F_2(y)$ are

$$p_1(x) = \int_{\mathbb{R}} p(x,y)dy, p_2(y) = \int_{\mathbb{R}} p(x,y)dx,$$

which are called marginal distribution functions, respectively.

Definition 2.4 (Expectation)**.** *In general case, we will use a special integral, called Riemann-Stieltjes integral, to define it.*

Let the distribution function of the random vector $(\xi_1, ..., \xi_r)$ is $F(x_1, ..., x_r)$. The expectation of $(\xi_1, ..., \xi_r)$ is $(\mathsf{E}\xi_1, ..., \mathsf{E}\xi_r)$, where

$$\mathsf{E}\xi_t = \int_{\mathbb{R}^r} x_i dF(x_1, ..., x_n) = \int_{\mathbb{R}} x_i dF_i(x_i),$$

where F_i is the distribution function of ξ_i.

So in the case of discrete, $\mathsf{E}(\xi_i) = \sum_i x_i p(x_i)$; in the continuous case, $\mathsf{E}(\xi_i) = \int_{\mathbb{R}} x_i p_i(x_i) dx_i$.

Remark 2.5. *This definition is actually from the law of the unconscious statistician:*

$$\mathsf{E}g(\xi_1, ..., \xi_d) = \int_{\mathbb{R}^d} g(x_1, ..., x_d) dF(x_1, ..., x_d).$$

Definition 2.6 (Variance and Covariance)**.** *Consider a random vector $\boldsymbol{\xi} = (\xi_1, ..., \xi_r)$, we define its Covariance matrix as*

$$\mathsf{Cov}(\boldsymbol{\xi}) = \begin{pmatrix} \mathsf{Var}\xi_1 & \mathsf{Cov}(\xi_1, \xi_2) & \cdots & \mathsf{Cov}(\xi_1, \xi_r) \\ \mathsf{Cov}(\xi_2, \xi_1) & \mathsf{Var}\xi_2 & \cdots & \mathsf{Cov}(\xi_2, \xi_r) \\ \vdots & \vdots & & \vdots \\ \mathsf{Cov}(\xi_r, \xi_1) & \mathsf{Cov}(\xi_r, \xi_2) & \cdots & \mathsf{Var}\xi_r \end{pmatrix},$$

where $\mathsf{Var}\xi_i = \mathsf{E}(\xi_i - \mathsf{E}\xi_i)^2$ is the variance of ξ_i and $\mathsf{Cov}(\xi_i, \xi_j) = \mathsf{E}(\xi_i - \mathsf{E}\xi_i)(\xi_j - \mathsf{E}\xi_j)$ is the covariance of ξ_i, ξ_j. So $\mathsf{Cov}(\xi_i, \xi_i) = \mathsf{Var}\xi_i$ and $\mathsf{Cov}(\boldsymbol{\xi}) = (\mathsf{Cov}(\xi_i, \xi_j))_{r \times r}$.

Remark 2.7. *(1) So we can use the law of the unconscious statistician to calculate them;*

(2) The matrix $\mathsf{Cov}(\boldsymbol{\xi})$ is non-negative definited since for all $t_j \in \mathbb{R}$ we have

$$\sum_{j,k} \mathsf{Cov}(\xi_j, \xi_k) = \mathsf{E} \left(\sum_j t_j (\xi_j - \mathsf{E}\xi_j) \right)^2 \geq 0.$$

3 Main Results and Their Proofs

3.1 Farlie-Gumbel-Morgenstern (FGM) copula

We will discover how to construct from two single distributions to a bivariant distribution and discover its properties such as FGM copula in [1],[8] and so on.

Definition 3.1 (Farlie-Gumbel-Morgenstern (FGM) copula)**.** *For two distributions $U \sim F_1(u), V \sim F_2(v)$, we can get the FGM copula as*

$$C(U,V) = UV(1+\alpha(1-U)(1-V)),$$

where $\alpha \in [-1,1]$. if α is greater than zero, the FGM copula provides positive dependence; if α is smaller than zero, it returns negative dependence; when α is zero, it reduces to the independence copula.

Remark 3.2. *We can see that the marginals of FGM copula is F_1, F_2, respectively. Moreover, in [2], we have the more generalized version of FGM copula.*

Now we start. Let $G(x,y) = F_1(x)F_2(y)(1+\alpha(1-F_1(x))(1-F_2(y)))$, then its probability density function is $g(x,y) = \partial^2_{x,y} G = f_1(x)f_2(y)(1+\alpha(2F_1(x)-1)(2F_2(y)-1))$. So we have

$$\begin{aligned}
\mathsf{Cov}(x,y) &= \int_{\mathbb{R}^2} (x-\mathsf{E}x)(y-\mathsf{E}y)g(x,y)dxdy \\
&= \left(\int_{\mathbb{R}} (x-\mathsf{E}x)f_1(x)dx\right)\left(\int_{\mathbb{R}} (y-\mathsf{E}y)f_2(y)dy\right) \\
&+ \alpha\left(\int_{\mathbb{R}} (x-\mathsf{E}x)f_1(x)(2F_1(x)-1)dx\right)\left(\int_{\mathbb{R}} (y-\mathsf{E}y)f_2(y)(2F_2(y)-1)dx\right) \\
&= \alpha\left(\int_{\mathbb{R}} (x-\mathsf{E}x)f_1(x)(2F_1(x)-1)dx\right)\left(\int_{\mathbb{R}} (y-\mathsf{E}y)f_2(y)(2F_2(y)-1)dx\right) \\
&= \alpha\left(\int_{\mathbb{R}} xf_1(x)(2F_1(x)-1)dx\right)\left(\int_{\mathbb{R}} yf_2(y)(2F_2(y)-1)dx\right)
\end{aligned}$$

where the last step is because $\int_{\mathbb{R}}(2F_i(z)-1)f_i(z)dz = \int_0^1(2u-1)du = 0$. Also we have

$$\begin{aligned}
\rho(x,y) = \frac{\mathsf{Cov}(x,y)}{\mathsf{Var}(x)\mathsf{Var}(y)} &= \frac{\alpha}{\mathsf{Var}(x)\mathsf{Var}(y)} \\
&\times \left(\int_{\mathbb{R}} xf_1(x)(2F_1(x)-1)dx\right)\left(\int_{\mathbb{R}} yf_2(y)(2F_2(y)-1)dx\right).
\end{aligned}$$

Moreover we can find the upper bound of its correlation coefficient. We have

$$\begin{aligned}
&\left(\int_{\mathbb{R}} x(2F(x)-1)f(x)dx\right)^2 = \left(\int_{\mathbb{R}} (x-\mathsf{E}x)(2F(x)-1)f(x)dx\right)^2 \\
&\leq \left(\int_{\mathbb{R}} (x-\mathsf{E}x)^2 f(x)dx\right)\left(\int_{\mathbb{R}} (2F(x)-1)^2 f(x)dx\right) = \frac{\mathsf{Var}(x)^2}{3},
\end{aligned}$$

where the last step is because $\int_{\mathbb{R}}(2F(z)-1)^2 f(z)dz = \int_0^1 (2u-1)^2 du = \frac{1}{3}$. So

$$\rho(x,y) = \frac{\alpha}{\mathsf{Var}(x)\mathsf{Var}(y)} \left(\int_{\mathbb{R}} x f_1(x)(2F_1(x)-1)dx \right) \times \left(\int_{\mathbb{R}} y f_2(y)(2F_2(y)-1)dx \right) \leq \frac{\alpha}{3},$$

well done.

3.2 Recurrence & Transience in Random Walk

Random walks are one of the most interesting and important example in probability theory which can show what difference between 2-dimension and more higher dimension. The motivation comes from observations of various random motions in physical and biological sciences. For now we will show the much difference between the simple symmetric random walk on $\mathbb{Z}^2$ and on $\mathbb{Z}^d (d \geq 3)$.

Definition 3.3 (Random Walk on $\mathbb{Z}^d$). *Let $X_1, ...$ be a sequence of $\mathbb{Z}^d$-valued independent and identically distributed random variables. A **random walk** started at $z \in \mathbb{Z}^d$ is the sequence $(S_n)_{n\geq 0}$ on $\mathbb{Z}^d$ where $S_0 = z$ and $S_n = S_{n-1} + X_n, n \geq 1$*

Definition 3.4 (Recurrence & Transience). *We say that a random walk is **recurrent** if it visits its starting position infinitely often with probability one and **transient** if it visits its starting position finitely often with probability one.*

This we refer [3]. Now we let $N = \sum_{n\geq 0} \mathbb{1}_{\{S_n = S_0\}}$. So recurrent means $\mathsf{P}(N = \infty) = 1$ and transience means $\mathsf{P}(N < \infty) = 1$. Let $\tau = \inf\{n \geq 1 : X_1 + \cdots + X_n = 0\}$ denote the first time the walk is back to the starting point. For now we will state some conclusion without proofs.

Lemma 3.5 (Either Recurrent or Transient)**.** *For each $n \geq 1$,*

$$\mathsf{P}(N = n) = \mathsf{P}(\tau = \infty)\mathsf{P}(\tau < \infty)^{n-1}.$$

In particular, every random walk is either recurrent or transient.

Lemma 3.6. *A random walk is transient if $\mathsf{E}N < \infty$ and recurrent if $\mathsf{E}N = \infty$.*

Proof. We can prove that $\mathsf{E}N = \frac{1}{\mathsf{P}(\tau=\infty)}$. ◻

Definition 3.7 (Simple Symmetric Random Walk)**.** *A random walk over $\mathbb{Z}^d$ is called Simple Symmetric Random Walk (SSRW) if all X_k take values over $\{\pm e_1, ..., \pm e_d\}$ and $P(X_k = (\pm)e_i) = \frac{1}{2d}$.*

Lemma 3.8 (Expectation Formula)**.** *Let $\phi(k) = \mathsf{E}(e^{ikX_1}) = \mathsf{E}(e^{ikX_i})$, then*

$$\mathsf{E}N = \lim_{t\uparrow 1} \int_{[-\pi,\pi]^d} \frac{1}{1 - t\phi(k)} \frac{dk}{(2\pi)^d}.$$

Sketch of the Proof. The proof based on

$$\mathbb{1}_{\{S_n=0\}} = \int_{[-\pi,\pi]^d} e^{ikS_n} \frac{dk}{(2\pi)^d},$$

then we have

$$\mathsf{P}(S_n = 0) = \int_{[-\pi,\pi]^d} \mathsf{E}(e^{ikS_n}) \frac{dk}{(2\pi)^d}.$$

Since S_n is the sum $X_1 + \cdots + X_n$, we find that

$$\mathsf{E}(e^{ikS_n}) = \mathsf{E}(e^{ikX_1}) \cdots \mathsf{E}(e^{ikX_n}) = \phi(k)^n.$$

So we have

$$\sum_{n=0}^{\infty} t^n \mathsf{P}(S_n = 0) = \int_{[-\pi,\pi]^d} \frac{1}{1 - t\phi(k)} \frac{dk}{(2\pi)^d}$$

and take $t \uparrow 1$ and we win. ◻

The following theorem is what we want and this observation allow us to characterize when the simple random walk is recurrent and when it is transient.

Theorem 3.9 (Recurrence/Transience of SSRW)**.** *The simple symmetric random walk on $\mathbb{Z}^d$ is recurrent in dimensions $d \leq 2$ and transient in dimensions $d > 2$.*

Proof. We have

$$\phi(k) = \frac{1}{2d} \sum_{j=1}^{d} (e^{ik_j} + e^{-ik_j}) = \frac{1}{d} \sum_{j=1}^{d} \cos(k_j).$$

By $1 - \cos x = 2\sin^2(x/2)$ inequalities $\frac{2x}{\pi} \leq \sin x \leq x$, we have

$$1 - t + 2t\frac{|k|^2}{\pi^2 d} \leq 1 - t\phi(k) \leq 1 - t + \frac{|k|^2}{\pi^2 d}.$$

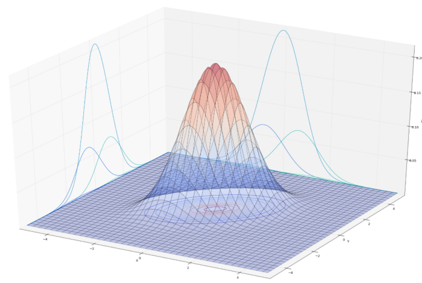
Taking the limit we find that the function $k \mapsto 1 - t\phi(k)$ is uniformly integrable around $k = 0$ if and only if the function $k \mapsto |k|^2$ is integrable, i.e.,

$$\mathsf{E}N < \infty \text{ if and only if } \int_{|k|<1} \frac{dk}{|k|^2} < \infty,$$

Since we find that the integral is finite if and only if $d > 2$, we win! ◻

3.3 Bivariate Normal Distribution

Bivariate normal distribution is one of the most important examples in probability and statistics and we will discover it here.



The probability density function of $\mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ is

$$p(x,y) = \frac{1}{2\pi\sigma_2\sigma_2(1-\rho^2)} \exp\left[-\frac{1}{2(1-\rho^2)}\times \left(\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right]$$

where $\sigma_i > 0 (i = 1,2), |\rho| < 1$. If we let

$$\boldsymbol{\mu} = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \boldsymbol{C} = \begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}$$

where $\sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2, \sigma_{ii} = \sigma_i^2$. Then we can rewrite it as

$$p(\boldsymbol{x}) = \frac{1}{2\pi(\det \boldsymbol{C})^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right].$$

Theorem 3.10 (Canonical Decomposition of Bivariate Normal Distribution)**.** *We have*

$$p(x,y) = \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right)\times\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{(y-(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)))^2}{2\sigma_1^2(1-\rho^2)}\right).$$

Proof. This is just a directly calculation and I will omit it. □

In this theorem we actually can rewrite it as:

$$\mathcal{N}_{x,y}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = \mathcal{N}_x(\mu_1, \sigma_1^2) \times \mathcal{N}_y(\mu_1 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2))$$

Proposition 3.11. *The marginals of bivariate normal distribution $\mathcal{N}_{x,y}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ is $\mathcal{N}_x(\mu_1, \sigma_1^2)$ and $\mathcal{N}_y(\mu_2, \sigma_2^2)$.*

Proof. Just use the theorem, we have

$$\begin{aligned}
p_1(x) = \int_{\mathbb{R}} p(x,y)dy &= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) \\
&\times \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{(y-(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)))^2}{2\sigma_1^2(1-\rho^2)}\right)dy \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right)
\end{aligned}$$

and another is similar. □

Remark 3.12. *So we see that the marginals of bivariate normal distribution are normal distributions, but the converse is not true. Actually, we let* $\phi(x) = \frac{1}{2\pi}e^{-x^2/2}, -\infty < x < \infty$ *and let*

$$g(x) = \begin{cases} \cos x, & |x| < \pi; \\ 0, & |x| \geq \pi. \end{cases}$$

Now we let $p(x,y) = \phi(x)\phi(y) + \frac{1}{2\pi}e^{-\pi^2}g(x)g(y)$*. For more counterexamples, we refer [7].*

For now we explore the meanings of μ_i, σ_i, ρ. By the proposition of marginals, we know that μ_i, σ_i^2 are expectations, repectively variances of its marginals. Then we will see what the ρ are.

Actually let $z = \frac{1}{\sqrt{1-\rho^2}}(\frac{x-\mu_1}{\sigma_1} - \rho\frac{y-\mu_2}{\sigma_2}), t = \frac{y-\mu_2}{\sigma_2}$, then we have

$$\begin{aligned}
\sigma_{12} &= \int_{\mathbb{R}^2}(x-\mu_1)(y-\mu_2)p(x,y)dxdy \\
&= \frac{1}{2\pi}\int_{\mathbb{R}^2}(\sigma_1\sigma_2\sqrt{1-\rho^2}tz + \rho\sigma_1\sigma_2 t^2)e^{-(t^2-z^2)/2}dzdt = \rho\sigma_1\sigma_2,
\end{aligned}$$

so $\sigma_{12} = \mathsf{Cov}(x,y) = \rho\sigma_1\sigma_2$ and $\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}$ is the correlation coefficient. So $\boldsymbol{C} = \mathsf{Cov}(\boldsymbol{\xi})$ where $\boldsymbol{\xi} \sim \mathcal{N}(\mu_i, \sigma_i, \rho)$.

Corollary 1. *At the case of bivariate normal distribution, uncorrelation if and only if independence.*

Proof. By the discuss above, independence if and only if $\rho = 0$ which is of course if and only if uncorrelation. □

Finally, we will consider how the distribution changes after act a linear transform.

Proposition 3.13. *Consider a matrix* $\boldsymbol{M}$ *and* $\boldsymbol{\xi} \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$*. For convenience we let* $\boldsymbol{\mu} = (\mu_1, \mu_2)^T, \boldsymbol{C} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$*. Then* $\boldsymbol{\eta} = \boldsymbol{M}\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{M}\boldsymbol{\mu}, \boldsymbol{M}\boldsymbol{C}\boldsymbol{M}^T)$*.*

Proof. Consider the characteristic function, for any real vector $\boldsymbol{t}$, we have

$$\begin{aligned}
f_{\boldsymbol{\eta}}(\boldsymbol{t}) &= \mathsf{E}e^{i\boldsymbol{t}^T\boldsymbol{M}\boldsymbol{\xi}} = \mathsf{E}e^{i(\boldsymbol{M}^T\boldsymbol{t})^T\boldsymbol{\xi}} \\
&= \exp\left(i(\boldsymbol{M}\boldsymbol{\mu})^T\boldsymbol{t} - \frac{1}{2}\boldsymbol{t}^T(\boldsymbol{M}\boldsymbol{C}\boldsymbol{M}^T)\boldsymbol{t}\right),
\end{aligned}$$

so $\boldsymbol{\eta} = \boldsymbol{M}\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{M}\boldsymbol{\mu}, \boldsymbol{M}\boldsymbol{C}\boldsymbol{M}^T)$. □

4 Conclusions

For now we have discoverd the properties of FGM copula, the difference between 2-dimension and more higher dimension distribution by exploring the simple symmetric random walk and some properties about the bivariate normal distribution.

5 Symbols and Notations

Table 1: Symbols

Symbol	meaning	Symbol	meaning
P	Probability Measure	p	Probability Density Function
E	Expectation	Var	Variance
Cov	Covariance	ρ	Correlation Coefficient
$\mathbb{1}$	Indicator Function	$\mathcal{N}$	Normal Distribution

References

[1] Alessandro Barbiero. A bivariate geometric distribution allowing for positive or negative correlation. *Communications in Statistics-Theory and Methods*, 48(11):2842–2861, 2019.

[2] Hakim Bekrizadeh, Gholam Ali Parham, and Mohamad Reza Zadkarmi. The new generalization of farlie–gumbel–morgenstern copulas. *Applied Mathematical Sciences*, 6(71):3527–3533, 2012.

[3] Marek Biskup. *PCMI Undergraduate Summer School Lecture notes*. Not Published, 2007.

[4] D. J. G. Farlie. The performance of some correlation coefficients for a general bivariate distribution. *Biometrika*, 47(3/4):307–323, 1960.

[5] Albert N. Shiryaev. *Probability-1*. Springer, New York, NY, 2016.

[6] Martin J. Sklar. Fonctions de repartition a n dimensions et leurs marges. 1959.

[7] Jordan Stoyanov. *Counterexamples in Probability*. Dover Publications Inc., 2014.

[8] Tsutomu T Takeuchi. Constructing a bivariate distribution function with given marginals and correlation: application to the galaxy luminosity function. *Monthly Notices of the Royal Astronomical Society*, 406(3):1830–1840, 2010.

[9] Susanne Trick, Frank Jäkel, and Constantin A Rothkopf. A bivariate beta distribution with arbitrary beta marginals and its generalization to a correlated dirichlet distribution. *arXiv preprint arXiv:2104.08069*, 2021.

[10] Christopher Withers and Saralees Nadarajah. An identity for the bivariate normal. *The Mathematical Scientist*, 34, 01 2009.

[11] Li Xianping. *Fundamentals of Probobility, 3rd Edition.* Higher Education Press, 2010.