“This is my design.”

PAVLOV, POISSON AND PARZEN

C.D.R.

Contents

Chapter 1

Count Distributions

1.1 Discrete Random Variables

For the below, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}), \#)$ be the discrete space, integers equipped with the counting measure.

Definition 1.1.1 (Discrete Random Variable)**.** A *discrete random variable* is a function $X : \Omega \to \mathbb{Z}$ such that for each $Z \in \mathcal{P}(\mathbb{Z})$, $X^{-1}(Z) \in \mathcal{F}$.

Definition 1.1.2 (Discrete Probability Distribution)**.** The *distribution* of X is the *pushforward measure* $X_*\mathbb{P} : \mathcal{P}(\mathbb{Z}) \to [0, 1]$ given by:

$$X_*\mathbb{P}(Z) = \mathbb{P}(X^{-1}(Z)) = \mathbb{P}(X \in Z)$$

We say that $X \sim \mathcal{D}$ iff $\forall Z \in \mathcal{P}(Z)$, $X_*\mathbb{P}(Z) = \mathcal{D}(Z)$. However, we will almost always characterise distributions using probability mass functions.

1.1.1 Distribution Functions

To understand properly the probability mass function, we require notions of σ-finiteness and absolute continuity of measures.

Definition 1.1.3. A measure μ on a measurable space $(\Omega, \mathcal{F})$ is *σ-finite* if, for each $F \in \mathcal{F}$ with $\mu(F) = \infty$, F can be written as a countable union of sets in $\mathcal{F}$, each with finite measure under μ.

Definition 1.1.4. Given measures μ, ν on $(\Omega, \mathcal{F})$, we say that μ is *absolutely continuous* with respect to ν and write $\mu \ll \nu$, if for each $F \in \mathcal{F}$, $\nu(A) = 0 \Rightarrow \mu(A) = 0$.

Radon-Nikodym Theorem. *Suppose μ, ν are two σ-finite measures on $(\Omega, \mathcal{F})$ with $\mu \ll \nu$. Then there exists a unique*[1] *measurable function $f : \Omega \to [0, \infty)$ such that for each $F \in \mathcal{F}$,*

$$\mu(F) = \int_F f d\nu$$

f is called the Radon-Nikodym derivative and is denoted $\frac{d\mu}{d\nu}$.

Definition 1.1.5 (Probability Mass Function)**.** The *probability mass function* of a distribution is given by the Radon-Nikodym derivative of the pushforward measure with respect to the counting measure. Namely any measurable function $f : \mathbb{Z} \to [0, 1]$ such that $\forall Z \subseteq \mathbb{Z}$,

$$X_*\mathbb{P}(Z) = \int_Z f d\#$$

[1] ν-almost everywhere

Theorem 1.1.1. *Every probability distribution admits a probability mass function.*

Proof. First observe that, should the derivative exist its range must be limited from $[0, \infty)$ to the range of the pushforward measure $[0, 1]$ as the integral is summing values in the function range. To show existence, it suffices, by the Radon-Nikodym theorem, to show that every discrete distribution is absolutely continuous with respect to the counting measure. Suppose that $\#(Z) = 0$ then $Z = \varnothing$ and by definition of a measure, $X_*\mathbb{P}(\varnothing) = 0$. $\square$

Definition 1.1.6 (Cumulative Distribution Function)**.** The *cumulative distribution function* of a distribution $F : \mathbb{Z} \to [0, 1]$, with mass function f, is given by the distribution function of an interval:

$$F(x) = X_*\mathbb{P}([0, \max\{0, x\}]) = \int_{[0,\max\{0,x\}]} f d\# = \sum_{i=0}^{\max\{0,x\}} f(x)$$

We define a left inverse of the c.d.f. to generate count distribution variables from the uniform distribution. It is a consequence of Cantor's work that no right inverse, and hence inverse, exists.

Definition 1.1.7. The *inverse c.d.f.* $F^{-1} : [0, 1] \to \mathbb{Z}$ of the a distribution is given by:

$$F^{-1}(u) = \min\{k \in \mathbb{Z} : u \leq F(k)\}$$

1.1.2 Variable Spaces

In general, the underlying measure space $(\Omega, \mathcal{F}, \mathbb{P})$ of an experiment is rarely explicitly characterized or even characterizable and so we will denote the set of all random variables with discrete space domain as $\mathcal{L}$.

Definition 1.1.8 (The Distribution Space)**.** The *distribution space* of a discrete distribution $\mathcal{D}$ is given by the set of all discrete random variables with distribution $\mathcal{D}$, denoted:

$$\mathcal{L}(\mathcal{D}) = \{X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{Z}, \mathcal{P}(\mathbb{Z}), \#) : X \sim \mathcal{D}\} = \{X \in \mathcal{L} : X \sim \mathcal{D}\}$$

1.2 Uniform Draw Representation

For the below, we require the notion of a continuous random variable. The domain measure space is $([0, 1], L[0, 1], \lambda)$, the unit interval with the Lebesgue σ-algebra and the Lebesgue measure. The definition is analogous to the discrete case.

Definition 1.2.1. A random variable $U : \Omega \to [0, 1]$ is *uniformly distributed* if for each measurable subset $Z \subseteq [0, 1]$, $\mathbb{P}(U \in Z) = \lambda(Z)$.

Note that this implies that $\forall a, b, k \in [0, 1]$ with $a \leq b$,

$$\mathbb{P}(k \in [a, b]) = \mathbb{P}(k \in (a, b]) = \mathbb{P}(k \in [a, b)) = \mathbb{P}(k \in (a, b)) = b - a$$

Definition 1.2.2 (Uniform Draw Variable)**.** For a uniform random variable U, the associated *uniform draw variable* of a distribution with c.d.f. F is given by $F^{-1}(U)$.

Proposition 1.2.1. $F^{-1}(U)$ *has the distribution associated with* F.

Proof. It suffices to show that the variable has the density that defines the cumulative distribution function.

$$\begin{aligned}
\mathbb{P}(F^{-1}(U) = k) &= \mathbb{P}(\min\{k \in \mathbb{Z} : U \leq F(k)\} = k) \\
&= \mathbb{P}(F(k - 1) < U \leq F(k)) \\
&= \mathbb{P}(U \in (F(k - 1), F(k)]) = F(k) - F(k - 1) = f(k)
\end{aligned}$$

$\square$

Definition 1.2.3. A function $q : [0, 1] \to [0, 1]$ is called a *mean involution* if for any uniformly distributed random variable U, $\mathbb{P}(F^{-1}(q(U)) = k) = f(k)$.

1.3 Poisson Distribution

We define the Poisson distribution as the limit of a binomial distribution.

1.3.1 Poisson Limit Theorem

Definition 1.3.1. Suppose $X \in \mathcal{L}$. We say that $X \sim \text{Bin}(n, p)$ with $n \in \mathbb{Z}$ and $p \in [0, 1]$ if:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Suppose that $X \in \mathcal{L}$ has a binomial distribution with n trials and $\mathbb{E}(X) = \lambda \in \mathbb{R}_{\geq 0} < n$. Then the probability of success of each trail is λ/n and as the number of trials tends to infinity, the random variable X becomes a infinite number of infinitesimally probable trials in some fixed time period, whose number of successes is expected to be λ. This is the *Poisson distribution* and its p.m.f. is given by the limit of the binomial as $n \to \infty$.

Proposition 1.3.1.

$$\lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

Definition 1.3.2 (Poisson Distribution)**.** Suppose $X \in \mathcal{L}$. We say that $X \sim \mathcal{P}(\lambda)$ with $\lambda \in \mathbb{R}_{\geq 0}$ if:

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

1.4 Poisson Shocks

1.4.1 Common Shock Model

1.4.2 Noble Shock Model

Definition 1.4.1. Suppose that $H \sim \mathcal{P}(\lambda_H)$, $A \sim \mathcal{P}(\lambda_A)$ are independent random variables and let $F^{-1}(U), F^{-1}(q(U)) \sim \mathcal{P}(\lambda_X)$ be uniform draw variables, independent of H and A. The common shock variable is then given by:

$$H + F^{-1}(U) \cap A + F^{-1}(q(U))$$

Proposition 1.4.1. *dd*

Proof.

$$\begin{aligned}
&\mathbb{P}(H + F^{-1}(U) \cap A + F^{-1}(q(U))) \\
&= \sum_{k=0}^{i} \mathbb{P}((H = k) \cap (F^{-1}(U) = i - k) \cap (A + F^{-1}(q(U)) = j)) \\
&= \sum_{l=0}^{j} \sum_{k=0}^{i} \mathbb{P}((H = k) \cap (F^{-1}(U) = i - k) \cap (A = l) \cap (F^{-1}(q(U)) = j - l)) \\
&= \sum_{l=0}^{j} \sum_{k=0}^{i} \mathbb{P}(H = k)\mathbb{P}(A = l)\mathbb{P}((F^{-1}(U) = i - k) \cap (F^{-1}(q(U)) = j - l))
\end{aligned}$$

□

Chapter 2

The Beautiful Game

2.1 Random Variables *or Measurable Functions*

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,1]}, \mathbb{P})$ be a filtered probability space and let the goals scored by the home team and away team respectively at time $t \in [0,1]$ be $H_t, A_t : \Omega \to \mathbb{N}$. Note that time t is given by the proportion of time elapsed so that for example: half time, or 45 minutes, is given by $t = 0.5$. We assume the goal variables are *homogeneous Poisson counting processes.*

Definition 2.1.1. A *homogeneous Poisson counting process* $\{N_t : t \geq 0\}$ is a stochastic process with

i. $N_t \geq 0$

ii. independent increments

iii. $N_t \sim \lambda_N t$

We will typically adopt a matrix notation for the *joint distributions* of H_1 and A_1 so that:

$$\alpha_{ij} = \mathbb{P}((H_1 = i) \cap (A_1 = j))$$

2.1.1 Empirical Probabilities

The aggregation of information and expertise can be found in the odds offered on betting exchanges for various events including, and particularly, football. The definition below is included for reference as it may feature in a future analysis of efficient markets.

Definition 2.1.2. Consider a bid-ask market M and suppose that b is the bid price and a the ask. Then the quoted spread Q is given by: $Q(M) = 2(a - b)/(a + b)$.

Efficient Market Hypothesis. *Let $h, a, d : [0,1] \to \mathbb{R}$ be the average of the back and lay odds, on a home win, away win and draw outcomes. Then,*

$$\textit{i.}\quad \mathbb{P}(H_1 > A_1) = \frac{1}{h} \qquad \textit{ii.}\quad \mathbb{P}(H_1 = A_1) = \frac{1}{d} \qquad \textit{iii.}\quad \mathbb{P}(A_1 > H_1) = \frac{1}{a}$$

2.2 Distributions *or Measures of Measurable Preimages*

We describe the joint distributions of the variables, in both dependent and independent contexts, under the assumptions that they follow Poisson distributions.

2.2.1 Independent Poisson

We first the consider the independent case characterised by the following equivalent statements:

i. $\mathbb{P}((H_1 = i) \cap (A_1 = j)) = \mathbb{P}(H_1 = i)\mathbb{P}(A_1 = j)$ ii. $\alpha_{ij} = \alpha_{i0}\alpha_{0j}$

Joint Distribution Matrix (IP1)**.**

$$u : \mathbb{R}^2_{\geq 0} \to M_n(\mathbb{R}) \qquad u_{ij}(\lambda_H, \lambda_A) = e^{-(\lambda_H+\lambda_A)}\frac{\lambda_H^i \lambda_A^j}{i!j!}$$

Derivation. Trivial from the Poisson p.d.f. and independence assumption. □

2.2.2 Dependent Poisson (Bivariate)

To consider the bivariate case, we make use of the *decomposability* of distributions.

Definition 2.2.1. A distribution $\mathcal{D}$ is *decomposable* if, for any random variable Z, there exist two independent, identically distributed variables X and Y such that $Z = X + Y$.

Theorem 2.2.1. *The Poisson distribution is decomposable.*

Therefore,

Joint Distribution Matrix (IP1)**.**

$$u : \mathbb{R}^2_{\geq 0} \to M_n(\mathbb{R}) \qquad u_{ij}(\lambda_H, \lambda_A) = e^{-(\lambda_H+\lambda_A)}\frac{\lambda_H^i \lambda_A^j}{i!j!}$$

Derivation. Trivial from the Poisson p.d.f. and independence assumption. □

Chapter 3

3.1 Introduction

Under the assumption that the full-time result market is efficient, we attempt to infer from the odds, the parameters of independent and bivariate Poisson distributions. These distributions are then used to find value in side markets with less liquidity and, hopefully, less liquidity.

3.1.1 Variables

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,1]}, \mathbb{P})$ be a filtered probability space and let the goals scored by the home team and away team respectively at time $t \in [0, 1]$ be $H_t, A_t : \Omega \to \mathbb{N}$. Note that time t is given by the proportion of time elapsed so that for example: half time, or 45 minutes, is given by $t = 0.5$. We assume the goal variables are *homogeneous Poisson counting processes.*

Definition 3.1.1. A *homogeneous Poisson counting process* $\{N_t : t \geq 0\}$ is a stochastic process with:

i. $N_t \geq 0$

ii. independent increments

iii. $N_t \sim \lambda_N t$

Definition 3.1.2. A *model* is a matrix of the form $\alpha_{ij} = \mathbb{P}((H_1 = i) \cap (A_1 = j))$.

3.1.2 Exchange Odds

Definition 3.1.3. Consider a bid-ask market M and suppose that b is the bid price and a the ask. Then the quoted spread Q is given by: $Q(M) = 2(a - b)/(a + b)$.

Efficient Market Hypothesis. *Let $h, a, d : [0, 1] \to \mathbb{R}$ be the average of the back and lay odds, on a home win, away win and draw outcomes. Then,*

$$i.\quad \mathbb{P}(H_1 > A_1) = \frac{1}{h} \qquad ii.\quad \mathbb{P}(H_1 = A_1) = \frac{1}{d} \qquad iii.\quad \mathbb{P}(A_1 > H_1) = \frac{1}{a}$$

3.2 Independent Poisson Model

3.2.1 Distribution

Joint Distribution Matrix (Independent)**.**

$$\alpha_{ij} : \mathbb{R}^2_{\geq 0} \to M_n(\mathbb{R}) \qquad\qquad \alpha_{ij} = e^{-(\lambda_H + \lambda_A)} \frac{\lambda_H^i \lambda_A^j}{i!j!}$$

Derivation.

$$\begin{aligned}\alpha_{ij} &= \mathbb{P}((H_1 = i) \cap (A_1 = j))\\ &= \mathbb{P}(H_1 = i)\mathbb{P}(A_1 = j)\\ &= e^{-\lambda_H}\frac{\lambda_H^i}{i!}e^{-\lambda_A}\frac{\lambda_A^j}{j!}\\ &= e^{-(\lambda_H+\lambda_A)}\frac{\lambda_H^i\lambda_A^j}{i!j!}\end{aligned}$$

□

3.2.2 Full-time Result Probabilities

Home Win

$$\begin{aligned}\mathbb{P}(H_1 > A_1) &= e^{-(\lambda_H+\lambda_A)}\sum_{i>j\geq 0}\frac{\lambda_H^i\lambda_A^j}{i!j!}\\ &= e^{-(\lambda_H+\lambda_A)}\Bigg(\lambda_H + \frac{\lambda_H^2\lambda_A}{2} + \frac{\lambda_H^2}{2} + \frac{\lambda_H^3\lambda_A^2}{12} + \frac{\lambda_H^3\lambda_A}{6} + \frac{\lambda_H^3}{6} + \frac{\lambda_H^4\lambda_A^3}{144} + \frac{\lambda_H^4\lambda_A^2}{48}\\ &\quad + \frac{\lambda_H^4\lambda_A}{24} + \frac{\lambda_H^4}{24} + \frac{\lambda_H^5\lambda_A^4}{2880} + \frac{\lambda_H^5\lambda_A^3}{720} + \frac{\lambda_H^4\lambda_A^2}{240} + \frac{\lambda_H^4\lambda_A}{120} + \frac{\lambda_H^5}{120} + \dots\Bigg)\end{aligned}$$

Draw

$$\begin{aligned}\mathbb{P}(H_1 = A_1) &= e^{-(\lambda_H+\lambda_A)}\sum_{i\geq 0}\frac{(\lambda_H\lambda_A)^i}{(i!)^2}\\ &= e^{-(\lambda_H+\lambda_A)}\left(1 + \lambda_H\lambda_A + \frac{(\lambda_H\lambda_A)^2}{4} + \frac{(\lambda_H\lambda_A)^3}{36} + \frac{(\lambda_H\lambda_A)^4}{576} + \frac{(\lambda_H\lambda_A)^5}{14400} + \dots\right)\end{aligned}$$

3.2.3 Away Win

$$\begin{aligned}\mathbb{P}(A_1 > H_1) &= e^{-(\lambda_A+\lambda_H)}\sum_{j>i\geq 0}\frac{\lambda_H^i\lambda_A^j}{i!j!}\\ &= e^{-(\lambda_A+\lambda_H)}\Bigg(\lambda_A + \frac{\lambda_A^2\lambda_H}{2} + \frac{\lambda_A^2}{2} + \frac{\lambda_A^3\lambda_H^2}{12} + \frac{\lambda_A^3\lambda_H}{6} + \frac{\lambda_A^3}{6} + \frac{\lambda_A^4\lambda_H^3}{144} + \frac{\lambda_A^4\lambda_H^2}{48}\\ &\quad + \frac{\lambda_A^4\lambda_H}{24} + \frac{\lambda_A^4}{24} + \frac{\lambda_A^5\lambda_H^4}{2880} + \frac{\lambda_A^5\lambda_H^3}{720} + \frac{\lambda_A^4\lambda_H^2}{240} + \frac{\lambda_A^4\lambda_H}{120} + \frac{\lambda_A^5}{120} + \dots\Bigg)\end{aligned}$$

3.3 Bivariate Poisson Model

3.3.1 Distribution

Joint Distribution Matrix (Independent)**.**

$$\alpha_{ij} : \mathbb{R}^3_{\geq 0} \to M_n(\mathbb{R}) \qquad\qquad \alpha_{ij} = e^{-(\lambda_H+\lambda_A)}\frac{\lambda_H^i\lambda_A^j}{i!j!}$$

Derivation.

$$
\begin{aligned}
\alpha_{ij} &= \mathbb{P}((H_1 + X_1 = i) \cap (A_1 + X_1 = j)) \\
&= \sum_{k=0}^{i} \mathbb{P}((H_1 = k) \cap (X_1 = i - k) \cap (A_1 + X_1 = j)) \\
&= \sum_{l=0}^{j} \sum_{k=0}^{i} \mathbb{P}((H_1 = k) \cap (X_1 = i - k) \cap (A_1 = l) \cap (X_1 = j - l)) \\
&= \sum_{l=0}^{j} \sum_{k=0}^{i} \mathbb{P}(H_1 = k)\mathbb{P}(X_1 = i - k)\mathbb{P}(A_1 = l)\mathbb{P}(X_1 = j - l) \\
&= \sum_{l=0}^{j} \mathbb{P}(A_1 = l) \sum_{k=0}^{i} \mathbb{P}(H_1 = k)\mathbb{P}(X_1 = i - k)\mathbb{P}(X_1 = j - l) \\
&= e^{-(\lambda_H + \lambda_A)} \frac{\lambda_H^i \lambda_A^j}{i!j!}
\end{aligned}
$$

□