Some Class

Daniel Yu

Contents

1 Poisson Distribution

Note. Balls to Bins
Given this setup: there are two bins with n number of balls and the balls are distributed uniformly at random across the 2 bins. Let's say that

$$P[N = 2] = \frac{1}{2}, P[N = 3] = \frac{1}{2}.$$

Let the random variables $X=$ number of balls in first bin and $Y=$ number of balls in the second bin. What is the distribution of X:

Proof. Consider the conditional probability

1. Distribution of X given $N = 2$: $X \sim Bin(2, .5)$
2. Distribution of X give $N = 3$: $X \sim Bin(3, .5)$

Use the law of total probability to construct the probability of X from the conditional probabilities of X:

$$P[X = 0] = P[N = 2] \cdot P[X = 0|N = 2] + P[N = 2] \cdot P[X = 0|N = 2] = \frac{1}{2} \cdot \frac{1}{2}^2 + \frac{1}{2}\frac{1}{2}^3 = \frac{3}{16}.$$

$$P[X = 1] = \frac{1}{2} \cdot \binom{2}{1} \frac{1}{2}^2 + \frac{1}{2} \cdot \binom{3}{1} \cdot \frac{1}{2}^3 = \frac{1}{4} + \frac{3}{16} = \frac{7}{16}.$$

$$P[X = 2] = \frac{5}{16}.$$

$$P[X = 3] = \frac{1}{16}.$$

Note that while the conditional probabilities are binomial, the combined probabilities is not binomial □

What about the distirbution of Y?

Proof. $Y \sim X$ □

Are X, Y independent?

Proof. Consider $P[X = 0, Y = 0] = 0$ and $P[X = 0] > 0$, $P[Y = 0] > 0$ □

Remark. However there is a special case where the above situation would have X, Y independent.

Definition 1. $N \sim Poisson(\lambda)$, $\lambda \geq 0$.

$$P[N = k] = \frac{e^{-\lambda}\lambda^k}{k!}.$$

Remark. Intuition

Let's repeat the experiment with this distribution for N.

$$
\begin{aligned}
P[X=r] &= \sum_{n=0}^{\infty} P[N=n]\cdot P[X=r|N=n] \\
&= \sum_{n=0}^{\infty} P[N=n]\cdot P[Bin(n,\frac{1}{2})=r] \\
&\text{if n > r, this is impossible} \\
&= \sum_{n=r}^{\infty} P[N=n]\cdot P[Bin(n,\frac{1}{2})=r] \\
&= \sum_{n=r}^{\infty} (\frac{e^{-\lambda}\lambda^n}{n!})[\binom{n}{r}\frac{1}{2}^n] \\
&= \frac{e^{-\lambda}}{r!}\sum_{n=r}^{\infty} -\frac{\frac{\lambda}{2}^n}{(n-r)!} \\
&\text{set m = n -r} \\
&= \frac{e^{-\lambda}}{r!}\sum_{m=0}^{\infty} \frac{\frac{\lambda}{2}^{m+r}}{m!} \\
&= \frac{e^{-\lambda}}{r!}\frac{\lambda}{2}^r\sum_{m=0}^{\infty} \frac{\frac{\lambda}{2}^m}{m!} \\
&\text{since } e^x = \sum_{m=0}^{\infty}\frac{x^m}{m!} \\
&= \frac{e^{-\frac{\lambda}{2}}\cdot\left(\frac{\lambda}{2}^n\right)}{r!} \\
&= \text{ Poisson}(\frac{\lambda}{2}).
\end{aligned}
$$

Similarly, $Y \sim Pos(\frac{\lambda}{2})$. Now, the claim is that for any s, r, $P[X=s, Y=r] = P[X=s]\cdot P[Y=r]$ that is X, Y independent.

$$
\begin{aligned}
P[X=s,Y=r] &= \frac{e^{-\frac{\lambda}{2}}\cdot\left(\frac{\lambda}{2}^n\right)}{r!}\cdot\frac{e^{-\frac{\lambda}{2}}\cdot\left(\frac{\lambda}{2}^n\right)}{s!} \\
&= \frac{e^{-\frac{\lambda}{2}}\cdot\left(\frac{\lambda}{2}^n\right)}{s!r!} \\
P[X=s,Y=r] &= P[N=r+s]\cdot P[X=s,Y=r|N=r+s] \\
&= \frac{e^{-\frac{\lambda}{2}}\cdot\left(\frac{\lambda}{2}^n\right)}{(r+s)!}\cdot P[X=s,Y=r|N=r+s] \\
&= \frac{e^{-\frac{\lambda}{2}}\cdot\left(\frac{\lambda}{2}^n\right)}{(r+s)!}\cdot P[X=s|N=r+s] \\
&= \frac{e^{-\frac{\lambda}{2}}\cdot\left(\frac{\lambda}{2}^n\right)}{(r+s)!}\cdot P[bin(r+s,\frac{1}{2})=s] \\
&= \frac{e^{-\frac{\lambda}{2}}\cdot\left(\frac{\lambda}{2}^n\right)}{(r+s)!}\cdot[\binom{r+s}{s}\cdot\left(\frac{1}{2}\right)^{r+s}] \\
&= \frac{e^{-\lambda}\left(\frac{\lambda}{2}^{r+s}\right)}{r!s!}.
\end{aligned}
$$

2 Exponential Distribution

Definition 2. $X \sim \exp(\lambda)$ if the pdf of X,

$$f_x(t) = \begin{cases} \lambda e^{-\lambda t}, t \geq 0 \\ 0, t < 0 \end{cases}.$$

Let $W \sim \exp(\lambda)$ the waiting between trains.

Note. Example
I arrived at the platform. Let W' be the amount of time I waiting between trains. I arrive to the platform. Let's suppose that the wat has 2 values : Long/short. $S = 5, L = 15$. Suppose you are told that 5 seconds have elapsed since the arrival of the previous train. Then $w' \sim w - s$ conditioned upon $w > s$. Let's understand,

$$\begin{aligned} P[w' > t] &= P[W > t + s \mid W > s] \\ &= \frac{P[W > t + s \cap W > s]}{P[W > s]} \\ &= \frac{P[W > t + s]}{P[W > s]} \\ &= \frac{e^{-\lambda(t-s)}}{e^{-\lambda s}} \\ &= e^{\lambda t}. \end{aligned}$$

So $w' \sim \exp(\lambda)$

3 Indepence and Identically Distributed

i.e. (iid) Random Variables

Note. IID Important Example
Suppose $\{X_i\}_{i=1}^n$ is a sequence of i.i.d random variables s.t. $P[X_i = 1] = P[X_i = 2] = P[X_i = 3] = P[X_i = 4]$. Let $k_i(n) =\mid \{j : x_j = i\} \mid =$ number of times i comes out. What do you expect $k_i(n)$ to be if n is large?

$$k_i(n) \sim \frac{1}{4}n.$$

Hence, $P[k_1(n) = n] = \frac{1}{2}^n > 0$ (all 1s for each n rolls). As we will show later,

$$P[k_1(n) = \frac{n}{4}] \approx \frac{1}{\sqrt{\pi} n}.$$

Definition 3. Let $\{X_n\}$ be a sequence of random variables. We say that X_n **converges in probability** to constant c if $\forall \varepsilon > 0$,

$$\lim_{n \to \infty} P[\mid X_n - c \mid > \varepsilon] = 0.$$

Note. Example
$G_n = Geo(\frac{n-1}{n})$, $P[G_n = k] = (\frac{n-1}{n}) \cdot (\frac{1}{n})^{k-1}$ to show that $G_n \to 1$ in probability.

Proof.

$$\forall \varepsilon > 0, P[|G_n - 1| > \varepsilon].$$

is small. as events $\{\mid G_n - 1 \mid > \varepsilon\} \subseteq \{G_n \geq 2\}$ (by set theory). So

$$P[\mid G_n - 1 \mid > \varepsilon] \subseteq P[G_n \geq 2] = \sum_{k=2}^{\infty} \left(\frac{n-1}{n}\right) \left(\frac{1}{n}^{k-1}\right) = \frac{1}{n}.$$

$$\lim_{n\to\infty} P[\mid G_n - 1 \mid \geq \varepsilon] \subseteq \lim_{n\to\infty} \frac{1}{n} = 0.$$

□

Note. Example
U_n is a uniform RV on $[-\frac{1}{n}, \frac{1}{n}]$. Claim $U_n \to 0$ in probability.

Proof.

$$\begin{aligned} P[|U_n| > \varepsilon] \to 0 \Leftrightarrow P[\mid U_n \mid \leq \varepsilon] &= \int_{-\varepsilon}^{\varepsilon} f_{U_n}(t)dt \\ &= \int_{-\varepsilon}^{\varepsilon} (\frac{n}{2}) 1|_{t\in[-\frac{1}{n},\frac{1}{n}]} dt \\ &= 1. \end{aligned}$$

Why? if $n > \frac{1}{\varepsilon}$ then $[-\frac{1}{n}, \frac{1}{n}] \subseteq (-\varepsilon, \varepsilon)$. □

3.1 Fat Tail

incredibly rare event that affects the average outcome

Note. Example Let $P[R_n = 0] = 1 - \frac{1}{n}$, $P[R_n = n^2] = \frac{1}{n}$.

Proof.

$$P[|R_n| > \varepsilon] = P[R_n = n^2] = \frac{1}{n} \to 0.$$

but,

$$E[R_n] = 0 \cdot (1 - \frac{1}{n}) + n^2 \cdot \frac{1}{n} = n \to \infty.$$

□

Saint Petersburg Problem: https://www.wikiwand.com/en/articles/St.$_Petersburg_paradox$