

# 1 Probability Theory

## 1.1 Some ole' Binomial shits

### 1.1.1 Expected value and variance of Binomial

**Theorem 1.1.** Let  $X \approx \text{Bin}(n, p)$ , then the expected value and variance is given by

1.  $\mathbb{E}(X) = np$ ,
2.  $\text{Var}(X) = np(1 - p)$ .

### 1.1.2 Binomial approx to Poisson

**Theorem 1.2.** For  $\lambda = np$ ,

$$\lim_{n \rightarrow \infty} \binom{n}{k} p^k 1 - p^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

First, we need the following lemma

**Lemma 1.1.** Let  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$  be the Gamma function, then the following equation holds

$$\lim_{n \rightarrow \infty} \frac{\binom{n}{k}}{n^k} = \frac{1}{k!}.$$

*Proof of Lemma 1.1.* Note first that

$$\Gamma(k+1) = k\Gamma(k) = \dots = k!\Gamma(1) = k!.$$

and Stirling's Formula for Gamma function states that

$$\Gamma(x+1) = \sqrt{2\pi x} x^x e^{-x} \left(1 + \frac{1}{12x} \dots\right).$$

Then we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\binom{n}{k}}{n^k} &= \lim_{n \rightarrow \infty} \frac{n!}{n^k (n-k)! k!} \\ &= \lim_{n \rightarrow \infty} \frac{\Gamma(n+1)}{n^k \Gamma(n-k+1) \Gamma(k+1)} \\ &= \frac{1}{\Gamma(k+1)} \lim_{n \rightarrow \infty} \sqrt{\frac{n}{n-k}} \left(\frac{n}{e}\right)^n \left(\frac{e}{n-k}\right)^{n-k} \frac{1}{n^k}. \end{aligned}$$

Where we used the first order of Stirling's Formula for Gamma function in the last equality.

Next, we only need to prove the limit equals to 1:

$$\begin{aligned} \lim_{n \rightarrow \infty} \sqrt{\frac{n}{n-k}} \left(\frac{n}{e}\right)^n \left(\frac{e}{n-k}\right)^{n-k} \frac{1}{n^k} &= \lim_{n \rightarrow \infty} \sqrt{\frac{n}{n-k}} \frac{1}{e^k} \frac{(1-k/n)^k}{(1-k/n)^n} \\ &= 1, \end{aligned}$$

since

$$\begin{aligned} \sqrt{\frac{n}{n-k}} &\rightarrow 1, \\ (1-k/n)^k &\rightarrow 1, \\ \frac{(1-k/n)^k}{(1-k/n)^n} &\rightarrow e^{-k}, \end{aligned}$$

as  $n \rightarrow \infty$ . ■

Now we may prove the main theorem

*Proof of Theorem 1.2.* Let  $p = \gamma/n$  and fix  $k$ , we have

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{\binom{n}{k}}{n^k} &= \lim_{n \rightarrow \infty} \binom{n}{k} \left(\frac{\gamma}{k}\right)^k \left(1 - \frac{\gamma}{n}\right)^{n-k} \\
&= \lim_{n \rightarrow \infty} \frac{n^k}{k!} \left(\frac{\gamma}{k}\right)^k \left(1 - \frac{\gamma}{n}\right)^{n-k} \quad \text{by Lemma 1.1} \\
&= \frac{\gamma^k}{k!} \left[ \lim_{n \rightarrow \infty} \left(1 - \frac{\gamma}{n}\right)^n \right] \left[ \lim_{n \rightarrow \infty} \left(1 - \frac{\gamma}{n}\right)^k \right] \\
&= \frac{\gamma^k}{k!} e^{-\gamma} (1)
\end{aligned}$$

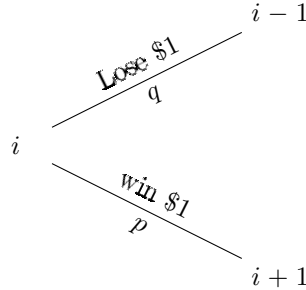
■

## 1.2 Gambler's Ruin

Let  $p$  is the probability of winning \$1 and  $q = 1 - p$  is the probability of losing \$1. Suppose that a person starts with \$ $i$  amount of money where  $0 < i < N$  and  $N \in \mathbb{N}$  is the number he would want to walk away with. The probability of him winning from this starting position is given by the following equation:

$$a_k = \begin{cases} \frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^N} & \text{where } p \neq q, \\ \frac{i}{N} & \text{where } p = q = \frac{1}{2}. \end{cases}$$

*Proof.* Note that



Then, we observe

$$\begin{aligned}
a_i &= pa_{i+1} + qa_{i-1} \\
\underbrace{(p+q)}_{=1} a_i &= pa_{i+1} + qa_{i-1} \\
a_{i+1} - a_i &= \frac{p}{q} (a_i - a_{i-1}).
\end{aligned}$$

Denote  $b_j := a_i - a_{i-1}$  and note that  $b_1 = a_1$  since  $a_0 = 0$  is the absorption edge, we deduce by recursion that

$$b_{i+1} = \left(\frac{p}{q}\right)^{i+1} a_1.$$

Moreover, note that

$$a_{i+1} - a_1 = \sum_{k=1}^{i+1} (a_k - a_{k-1})$$

$$\begin{aligned}
&= \sum_{k=1}^i \left(\frac{p}{q}\right)^k a_1. \\
\Rightarrow a_{i+1} &= a_1 \left(1 + \sum_{k=1}^i \left(\frac{p}{q}\right)^k\right) \\
&= a_1 \left(\sum_{k=0}^i \left(\frac{p}{q}\right)^k\right).
\end{aligned}$$

If  $p \neq q$ , by geometric series, we obtain

$$a_{i+1} = a_1 \frac{1 - \left(\frac{p}{q}\right)^{i+1}}{1 - \left(\frac{p}{q}\right)}.$$

If  $p = q = 0.5$ , then

$$a_{i+1} = a_1(i+1).$$

Now to solve for  $a_1$ , observe that since at absorption point  $a_N = 1$ , then letting  $i+1 = N$ , we get

$$1 = a_N = \begin{cases} \frac{1 - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)} & \text{where } p \neq q, \\ a_1 N & \text{where } p = q = \frac{1}{2}, \end{cases}$$

which implies

$$a_1 = \begin{cases} \frac{1 - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)} & \text{where } p \neq q, \\ \frac{1}{N} & \text{where } p = q = \frac{1}{2}. \end{cases}$$

Substituting this back, we obtained the desired equation. ■

## 2 Time Series

### 2.1 Stationary and Autocorrelation

**Definition 2.1.** *The autocovariance*

$$\gamma(s, t) := \mathbb{E}[(x_s - \mu_s)(x_t - \mu_t)],$$

and *cross-covar.*

$$\rho_{x,y}(s, t) := \mathbb{E}[(x_s - \mu_{xs})(y_t - \mu_{ts})].$$

**Definition 2.2.** *The autocorrelation function (ACF) is defined as*

$$\rho_x(s, t) := \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}},$$

**Notation:**  $\rho_x(t, t + h) \equiv \rho(h)$ .

**Definition 2.3** (Strictly stationary). *For all  $k \in \mathbb{N}$ ,  $t_k \in \mathbb{N}$  and  $c_k \in \mathbb{R}$ . Then the time series is strictly stationary if*

$$\mathbb{P}(x_{t_1} \leq c_1, \dots, x_{t_k} \leq c_k) = \mathbb{P}(x_{t_1+h} \leq c_1, \dots, x_{t_k+h} \leq c_k),$$

for given probability measure  $\mathbb{P}$  and constant  $h \in \mathbb{N}$ .

**Definition 2.4** (Weakly stationary). *The time series is strictly stationary if*

1.  $\mu_t = \mathbb{E}[x_t]$  is constant, and
2.  $\gamma(s, t)$  depends only on  $|s - t|$ .

**Lemma 2.1.** *Assume time series is weakly stationary. Using the notation  $\gamma(t, t + h) \equiv \gamma(h)$ , we have*

1.  $|\gamma(t)| \leq \gamma(0)$ ,
2.  $\gamma(h) = \gamma(-h)$ .

*Proof.* 1. By Cauchy-Schwarz inequality  $|\gamma(t, t + h)|^2 \leq \gamma(t, t)\gamma(t + h, t + h)$  by definition  $|\gamma(h)|^2 \leq \gamma(0)\gamma(0)$ .

2. From definition of covariance we have

$$\begin{aligned}\gamma(h) &= \gamma(t + h - t) \\ &= \mathbb{E}[(x_{t+h} - \mu)(x_t - \mu)] \\ &= \mathbb{E}[(x_t - \mu)(x_{t+h} - \mu)] \\ &= \gamma(t - (t + h)) = \gamma(-h)\end{aligned}$$

■

*Example 2.1.* Suppose  $w_t$  for all  $t \in \mathbb{N}$  is white noise process, i.e. iid  $w_t \approx N(0, \sigma_w^2)$ . Given two series

$$\begin{aligned}x_t &= w_t + w_{t-1}, \\ y_t &= w_t - w_{t-1}.\end{aligned}$$

Then for any  $h \in \mathbb{N}$ , we get the following for (cross-)covariance:

- 1.

$$\begin{aligned}\gamma_x(0) &= \mathbb{E}[(w_t + w_{t-1})^2] = \mathbb{E}[w_t^2] + \mathbb{E}[w_{t-1}^2] + 2 \underbrace{\mathbb{E}[w_t w_{t-1}]}_{=0} \\ &= 2\sigma_w^2, \\ \gamma_y(0) &= \mathbb{E}[(w_t - w_{t-1})^2] = \mathbb{E}[w_t^2] + \mathbb{E}[w_{t-1}^2] - 2 \underbrace{\mathbb{E}[w_t w_{t-1}]}_{=0} \\ &= 2\sigma_w^2.\end{aligned}$$

- 2.

$$\begin{aligned}\gamma_x(1) &= \mathbb{E}[(w_t + w_{t-1})(w_{t+1} + w_t)] \\ &= \mathbb{E}[w_t w_{t+1}] + \mathbb{E}[w_t^2] + \mathbb{E}[w_{t-1} w_{t+1}] + \mathbb{E}[w_{t-1} w_t] \\ &= \sigma_w = \gamma_x(-1), \\ \gamma_y(1) &= \mathbb{E}[w_t w_{t+1}] - \mathbb{E}[w_t^2] + \mathbb{E}[w_{t-1} w_{t+1}] - \mathbb{E}[w_{t-1} w_t] \\ &= -\sigma_w = \gamma_y(-1).\end{aligned}$$

3.  $\gamma_{xy}(0) = \mathbb{E}[w_t^2] - \mathbb{E}[w_{t-1}^2] = 0$  and  $\gamma_{xy}(1) = \text{cov}(x_{t+1}, y_t) = -\sigma_w^2$  and  $\gamma_{xy}(-1) = \text{cov}(x_t, y_{t-1}) = -\sigma_w^2$

#### 4. ACF

$$\rho_{xy}(h) = \begin{cases} 0 & h = 0, \\ 1/2 & h = 1, \\ -1/2 & h = -1, \\ 0 & |h| \geq 2 \end{cases} \quad (2.1)$$

Thus the joint time series is stationary.

**Definition 2.5** (Linear Process).  $\{x_t\}$  is linear process if

$$x_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}, \quad \sum_{j=-\infty}^{\infty} |\psi_j| < \infty,$$

where  $w_t$  is white noise.

**Definition 2.6** (Gaussian Process).  $\{x_t\}$  is Gaussian process if the vector  $x := (x_{t_1}, x_{t_2}, \dots, x_{t_n})' \in \mathbb{R}^n$  has a multivariate normal distribution with density function

$$f(x) = \frac{1}{(2\pi)^{n/2}} \det(\Gamma)^{-1/2} \exp\left(-\frac{1}{2}(x - \mu)' \Gamma^{-1} (x - \mu)\right),$$

where  $\mu \in \mathbb{R}^n$  and  $\Gamma := \text{var}(x) = \{\gamma(t_i, t_j); i, j = 1, \dots, n\}$ .

##### 2.1.1 ARIMA

ARMA( $p, q$ ):

$$x_t - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p} = w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}.$$

### 3 Autoregressive Model

**Definition 3.1** (AR( $p$ )). Let  $w_t$  be white noise, AR( $p$ ) is defined as

$$x_t - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p} = w_t.$$

**Definition 3.2** (Autoregressive Operator). Let  $B$  be time-lagged operator, i.e.  $Bx_t = x_{t-1}$ , the autoregressive operator is defined as

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p.$$

*Example 3.1.* Assuming  $\phi_j = \phi$  for all  $j \in \mathbb{N}$ , from AR(1)

$$\begin{aligned} x_t &= \phi x_{t-1} + w_t = \phi(\phi x_{t-2} + w_{t-1}) + w_t \\ &= \phi^2 x_{t-2} + \phi w_{t-1} + w_t \\ &\vdots \\ &= \phi^k x_{t-k} + \sum_{j=0}^{k-1} \phi^j w_{t-j}. \end{aligned}$$

Thus, if  $|\phi| < 1$  and  $x_t$  is stationary then for large  $k \rightarrow \infty$ , we have

$$x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j}.$$

Taking expectation value,

$$\mathbb{E}[x_t] = \sum_{j=0}^{\infty} \phi_j \mathbb{E}[w_t] = 0.$$

Moreover, for  $h \geq 0$  the autocovariance function is

$$\begin{aligned} \gamma(h) &= \mathbb{E} \left[ \left( \sum_{j=0}^{\infty} \phi^j w_{t+h-j} \right) \left( \sum_{k=0}^{\infty} \phi^k w_{t-k} \right) \right] \\ &= \mathbb{E} \left[ (w_{t+h-1} + \phi w_{t+h-2} + \cdots + \phi^h w_t + \phi^{h+1} w_{t-1} + \cdots) (w_t + \phi w_{t-1} + \cdots) \right] \\ &= \sigma_w^2 \sum_{j=0}^{\infty} \phi^{h+j} \phi^j = \sigma_w^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} \\ &= \frac{\sigma_w^2 \phi^h}{1 - \phi^2}. \end{aligned}$$

And ACF is

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h.$$

As well as

$$\rho(h) = \phi \rho(h-1).$$

**Definition 3.3.** *The moving average operator is defined as*

$$\theta(B) := 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q.$$

## 4 Operational Research

### 4.1 Safety Stock

**Theorem 4.1.** *Let the lead-time  $T$  be normally distributed, i.e.  $T \sim N(\mu_\ell, \sigma_\ell^2)$  and denote the set of demand of period  $i$  as  $\{D_i\}_{i=1}^T$  where  $D_i \sim N(\mu_d, \sigma_d^2)$  for each  $1 \leq i \leq T$ . Then the safety stock level,  $SS$  satisfies the following equation*

$$SS = z_\alpha \sqrt{\sigma_d^2 \mu_\ell + \mu_d^2 \sigma_\ell^2},$$

where  $z_\alpha$  is the Z-value of a desired  $\alpha$  which is chosen.

*Proof.* Denote the demand between within lead-time as  $D(T) := \sum_{i=1}^T D_i$ . The safety stock level is then set at

$$SS = z_\alpha \sqrt{\text{Var}(D(T))}.$$

Thus we need only to prove  $\text{Var}(D(T))$ .

From Towering-property, note that

$$\begin{aligned} \mathbb{E}[D(T)] &= \mathbb{E}[\mathbb{E}[D(T)|T]] \\ &= \mathbb{E}[T\mu_d] = \mu_d \mathbb{E}[T] = \mu_d \mu_\ell. \end{aligned}$$

Furthermore, again with Towering-property, we have

$$\begin{aligned} \mathbb{E}[D^2(T)] &= \mathbb{E}[\mathbb{E}[D^2(T)|T]] \\ &= \mathbb{E}[\text{Var}[D(T)|T] + (\mathbb{E}[D(T)|T])^2] \\ &= \mathbb{E}[T\sigma_d^2 + (T\mu_d)^2] \\ &= \sigma_d^2 \mathbb{E}[T] + \mu_d^2 \mathbb{E}[T^2] \\ &= \sigma_d^2 \mu_\ell + \mu_d^2 (\text{Var}(T) + (\mathbb{E}[T])^2) \\ &= \sigma_d^2 \mu_\ell + \mu_d^2 (\sigma_\ell^2 + \mu_\ell^2). \end{aligned}$$

Finally,

$$\begin{aligned} \text{Var}(D(T)) &= \mathbb{E}[D^2(T)] - (\mathbb{E}[D(T)])^2 \\ &= \sigma_d^2 \mu_\ell + \mu_d^2 (\sigma_\ell^2 + \mu_\ell^2) - (\mu_d \mu_\ell)^2 \\ &= \sigma_d^2 \mu_\ell + \mu_d^2 \sigma_\ell^2. \end{aligned}$$

■