Probability in Finance

Aditya Nair
Department of Electronics and Communication Engineering
IIIT Hyderabad
2020102022
aditya.nair@students.iiit.ac.in

Agrim Rawat
Department of Electronics and Communication Engineering
IIIT Hyderabad
2020102037
agrim.rawat@students.iiit.ac.in

Sankalp S. Bhat
Department of Electronics and Communication Engineering
IIIT Hyderabad
2020112018
sankalp.b@research.iiit.ac.in

Tejah S S
Department of Electronics and Communication Engineering
IIIT Hyderabad
2020112028
tejah.ss@research.iiit.ac.in

***Abstract*—This report is a brief overview on the involvement of probability and random processes in the field of finance.**

I. Introduction

In this section, we shall introduce the various terms used in the report. We refer to the asset as a random variable X, which is defined on the sample space

$$\Omega : \omega \to X(\omega), \omega \in \Omega$$

Here, all the unknown factors, such as the general economic and political climate, the past performance, natural disasters, etc. are included in the sample ω. We can broadly look at the sample in this way: we will have two sample points ω_1 and ω_2, which will denote the positive and the negative variation in the economic status of the world respectively. Then accordingly, if we assume that one has Rs.x_0 invested in an asset, then the holdings of that person according to the two sample points will be

$$X(\omega_1) = ax_0, a > 1, \text{ and}$$
$$X(\omega_2) = bx_0, b \in [0, 1)$$

This defines just one of the many cases involved in the stock market. There will be certain shares that will rise when sample point involved is ω_2 and decline when sample point involved is ω_1. It will be fair to assume that the sample space of our choice, ω is uncountable, considering the limitless possibilities in the economic world.

II. Asset Returns

We now define another random variable, ζ, which is defined as

$$\omega \in \Omega : \omega \to \zeta(\omega) = X(\omega) - x_0$$

This basically denotes the change in value of our asset at an instant, which we can clearly see will be denoted as a 'gain' if $\zeta > 0$, and a 'loss' if $\zeta < 0$.

To calculate how much return you get on your initial investment in an asset, we define another random variable as

$$R = \frac{\zeta(\omega)}{x_0}$$

Thus, from this we can basically write our current value, and predicted value of our asset as $x_0 = np$ and $X = nP$ respectively. Thus, we can represent our returns on the asset as

$$\zeta = \frac{P - p}{p} = \frac{P}{p} - 1$$

A. One-period and Multiperiod models

To define a random variable for an asset, we need to know its value at two separate durations of time t_0 and t_1, where, the value of the asset at time $t = t_0$ is constant and known to us, labeled x_0. We wish to find out the value of the asset at $t = t_1$, which can be depicted by a random variable.
Thus for a time interval T, we can define N periods over the interval $[0, T]$ such that $t_i - t_{i-1} = \frac{T}{N}$, where $(t_{i-1}, t_i]$ is an interval defined over $[0, T]$.

Hence, we can write the return of the asset over the i^{th} period (r_i)as follows:

$$r_i = \frac{P_i - P_{i-1}}{P_{i-1}}$$
$$r_i = \frac{P_i}{P_{i-1}} - 1$$

where, P_i is the price of the asset on the i^{th} date.

While the above equations define a multi-period model, a one-period model can be represented as the particular case of the above equation where $N = 1$.

B. Annualization

While the return of an asset is characterized by the change in the relative price over a particular period of time, and is infact actually a *rate of return* rather than a *return*, both the terms are used interchangeably in finance. The act of using standard rates of measure such as the change in percentage of asset value over a year is called *annualization*.

The *gross return* is defined as $1+r$, where r is the rate of return of an asset.

$$1 + r_i = 1 + \frac{P_i}{P_{i-1}} - 1 = \frac{P_i}{P_{i-1}}$$

C. Annualized rate of return

Let us assume multi-period model with N periods defined over a time interval of $T = 1$ year.

Then, we can calculate the gross return of the asset over N periods as follows:

$$\begin{aligned}\frac{P_N}{P_0} &= \frac{P_N}{P_{N-1}} \times \frac{P_{N-1}}{P_{N-2}} \times \cdots \frac{P_1}{P_0} \\ &= (1+r_N) \times (1+r_{N-1}) \times \cdots (1+r_1) \\ &= \prod_{i=1}^{N} (1+r_i)\end{aligned}$$

For N periods defined over a time interval of 1 year, We define a quantity *annualized rate of return* $AR^{(N)}$ as follows:

$$1 + AR^{(N)} = \frac{P_N}{P_0} = \prod_{i=1}^{N} (1+r_i).$$

For when the period is exactly a year, we find a fixed rate of return $AR^{(N)}$ such that

$$\begin{aligned}\frac{P_N}{P_0} &= (1+AR^{(N)})^N \\ \implies \log \frac{P_N}{P_0} &= N \cdot \log\left(1+AR^{(N)}\right) \\ \implies \log\left(1+AR^{(N)}\right) &= \frac{1}{N} \log\left(\prod_{i=1}^{N}(1+r_i)\right) \\ \implies & \boxed{\log\left(1+AR^{(N)}\right) = \frac{1}{N} \cdot \sum_{i=1}^{N} \log\left(1+r_i\right)}\end{aligned}$$

The above model is analogous to a compound interest investment model, as observed in many instances in real life.

D. Code

```python
import random

# Global Variables
good = ["A", "B", "C"]     # Good states
bad = ["D", "E", "F"]      # Bad states
sampleSpace = [good, bad] # All states

n = 100      # Number of shares
p0 = 5       # Initial price per share
x0 = n * p0 # Total initial price of shares

factor = random.random()  # Random number in [0, 1)
```

Fig. 1. Global Variables

```python
# Main function
def main():
    state = pdfX()
    P = X(state) / n
    Rx = (P - p0) / p0
    print("n      = {}".format(n))
    print()
    print("x0     = {}".format(x0))
    print("p0     = {}".format(p0))
    print()
    print("X(w)   = {}".format(X(state)))
    print("P      = {}".format(P))
    print()
    print("Del(w) = {}".format(Del(state)))
    print("R(w)   = {}".format(R(state)))
    print()
    print("R = (P - p0) / p0 = {}".format(Rx))
```

Fig. 2. Main Function

```python
# X : random variable representing the value of your holdings a year from now
def X(state):
    if state in good:  # Price of stock increases
        return (1 / factor) * x0
    elif state in bad: # Price of stock decreases
        return factor * x0


# Probability Distribution Function of X
def pdfX():
    p = 0.7 # Probability of good state

    k = random.random()
    if k < p:
        return random.choice(sampleSpace[0])
    else:
        return random.choice(sampleSpace[1])


# Del : random variable representing change in the value of your asset holdings
def Del(state):
    return X(state) - x0


# R : random variable representing the return on initial investment
def R(state):
    return Del(state) / x0
```

Fig. 3. Random Variable Definitions

III. Asset Risk

By asset risk, we mean the possibility of someone losing their net value in an asset, to the extent of losing everything related to that asset. We attempt to express this risk mathematically through a famous problem called the 'Gambler's ruin problem', which involves basic knowledge about random walks.

We first take an integer lattice on $\mathbb{R} = (-\infty, \infty)$, which we will denote by I. Now, take a particle at origin, which moves in the following fashion: the particle moves on the lines in I, it can only move to the left or the right. The probability of it moving to the left is taken as $p, p \in (0,1)$, and the probability of it moving to the right is $q = 1-p$. For convenience's sake, we assume that the particle has a speed of one step per unit time. Plotting the position of the particle as a function of time gives us the plot in "Fig. 5".

Let the independent random variable ξ_n denote the n^{th} step

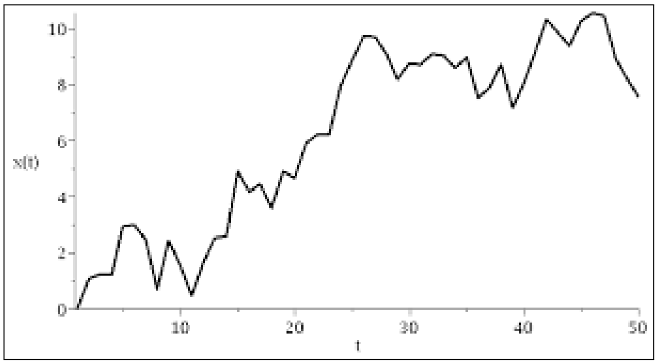

Fig. 4. A sample of Random Walk in one-dimension

of the particle defined as:

$$\xi_n = \begin{cases} +1 & \text{with probability } p \\ -1 & \text{with probability } q \end{cases}$$

If we denote the initial coordinates of the particles as X_0, then $X_n = X_0 + \sum_{i=1}^{n} \xi_i$. Thus we see that $\sum_{i=1}^{n} \xi_i$ is a sum of independent Bernoullian random variables. In relation with the 'gambler's ruin' problem, we look at the below problem.

In the sequence $X_n, n \geqslant 0$, will the particle ever hit a given point? In the 'gambler's ruin' problem , we take this given point as the point of bankruptcy of one of the players. Mathematically speaking, considering an interval $[0, z]$, where $z = x + y, x, y \geqslant 1$. Given that the particle starts from coordinate x, what is the probability that it will reach y? Solving equations which will not be shown here, we see that the probability that the particle will reach 0 before z when it starts from the point $i = [1, z-1]$ is

$$u_j = \frac{(\frac{q}{p})^i - (\frac{q}{p})^z}{1 - (\frac{q}{p})^z},\ 0 \leqslant i \leqslant z.$$

And the probability that the particle will reach z before 0 when it starts from the point i is

$$v_j = \begin{cases} \frac{1-(\frac{q}{p})^i}{1-(\frac{q}{p})^z} & p \neq q \\ \frac{i}{z} & p = q \end{cases}$$

A. Definition for risk of an asset

We define the risk associated with an asset, which has return R, as the variance $\boxed{\sigma^2(R)}$. We prefer to use the variance, instead of the standard deviation for mathematical convenience.

B. Riskless assets

As is obvious, if the variance $\sigma^2(R) = 0$, then the asset/security is termed as riskless. Suppose we take the rate of return for a year, r_a as a constant, then the value of asset X a year from now will be $X(\omega) = (1 + r_a)x_0\ \forall\ \omega \in \Omega$. Thus the return will be be

$$R(\omega) = \frac{X(\omega) - x_0}{x_0} = r_f$$

Thus the variance of return in terms of the expected return will be

$$\begin{aligned} \sigma^2(R) &= E\{(R - E(R))^2\} \\ &= E\{(r_f - r_f)^2\} \\ &= E\{0\} \\ \implies \boxed{\sigma^2(R) = 0} \end{aligned}$$

While talking about riskless assets here, we assume the following:

1) There are no transaction costs or taxes, and all the assets are perfectly divisible.
2) The investor believes that his actions cannot affect the probability distribution of returns on his available assets. So, if x_i is allocated from total X holdings to asset i, then $\{x_1, ..., x_n\}$ determines his holding's probability distribution.

IV. Portfolios

Suppose an investor has N given investment opportunities available to them, then, we can define an N tuple $(\alpha_1, \alpha_2, \ldots, \alpha_N)$ having N elements such that each α_i represents the proportion of total wealth invested in the i^{th} asset. It is elementary to observe that

$$\sum_{i=1}^{N} \alpha_i = 1$$
$$\text{where, } \ i \in \{1, 2, \ldots, N\}$$

We can also term the proportion of wealth invested in the i^{th} asset (α_i) as the investor's *portfolio weight*.

A. Diversification of Portfolios

We define the term 'diversification' of a portfolio for a new distribution of said portfolio, but with a lower risk. Mathematically speaking, if we have an N-tuple of portfolios $(\alpha_1, \alpha_2, \ldots, \alpha_N)$ with annual returns as R_i for a given security i, then the corresponding risk will be $\sigma^2(\sum_{i=1}^{N} \alpha_i R_i)$. To diversify this portfolio, we define a new tuple portfolio as $(\alpha'_1, \alpha'_2, \ldots, \alpha'_N)$, whose risk

$$\sigma^2\left(\sum_{i=1}^{N} \alpha'_i R_i\right) < \sigma^2\left(\sum_{i=1}^{N} \alpha_i R_i\right)$$

Example 1: Let us take a example to understand this better. Let us assume we have two assets with equal, positive risks. Let us also assume that the correlation coefficient between the returns of the two assets lies in the bounds of $[-1, 1]$. If we have the old portfolio as (α_1, α_2), let us assume the new portfolio to be $(\alpha, 1-\alpha)$. Thus, the risk related to the new portfolio will be

$$
\begin{aligned}
&\sigma^2(\alpha R_1 + (1-\alpha)R_2 \\
&= \sigma^2(R_1) + 2\alpha \cdot (1-\alpha)\left[Cov(R_1, R_2) - \sigma(R_1)\cdot\sigma(R_2)\right] \\
&< \sigma^2(R_1) = \sigma^2(\alpha_1 R_1 + \alpha_2 R_2)
\end{aligned}
$$

This can be interpreted as follows: when one has an opportunity to invest in 2 securities with equal risk, and the returns are either uncorrelated, or negatively correlated, then it is more beneficial/less risky to invest in both the securities, rather than investing in only one. To minimize the risk involved, we perform differentiation on the total variance term, and equate it to zero to get the value $\alpha = \frac{1}{2}$. Thus, this translates to investing equal capital in both the assets to ensure maximum diversification.

Example 2: Continuing from Example 1, we see if there is an allocation $(\alpha, 1-\alpha)$ where $0 < \alpha < 1$, that is least risky. That is, is there an $\alpha \in (0,1)$ that minimizes $V(\alpha)$ defined by

$$V(\alpha) = \sigma^2(\alpha R_1 + (1-\alpha)R_2)$$

From Example 1, we see that V is a quadratic function in α, with second derivative

$$V''(\alpha) = 2\sigma^2(R_1) + 2\sigma^2(R_2) - 4\rho_{12}\sigma(R_1)\sigma(R_2)$$

Assuming that $-1 \leq \rho_{12} < 1$ and $\sigma(R_1) = \sigma(R_2) > 0$, we have

$$V''(\alpha) > 2\sigma^2(R_1) + 2\sigma^2(R_2) - 4\sigma(R_1)\sigma(R_2)$$

$$> 2(\sigma(R_1) - \sigma(R_2))^2 = 0$$

Since $V''(\alpha) > 0$, to get the minimal α, say α^*, we solve the equation,

$$
\begin{aligned}
&V'(\alpha^*), \\
&= 2\alpha^*\sigma^2(R_1) - 2(1-\alpha^*)\sigma^2(R_2) + 2\rho_{12}\sigma(R_2)(1-2\alpha^*) = 0
\end{aligned}
$$

The above equation simplifies, after canceling $2\sigma^2(R_1)$ to,

$$\alpha^* - (1-\alpha^*) + \rho_{12}(1-2\alpha^*) = 0$$

Finally we can deduce that $\alpha^* = \frac{1}{2}$

What this means when we invest our fortune equally in each of two assets, we get the smallest return variance. In other words, if we have 2 assets with the same risk(same return variance), we get the most diversified portfolio. The allocation is independent of the strength of return correlation($\rho_{12} \neq 1$ between the two assets. The corresponding risk is

$$V(\alpha^*) = \frac{1}{2}(1-\rho_{12})\sigma^2(R_1) < \sigma^2(R_1)$$

since $\rho_{12} < 1$.
Note that if two assets are such that their returns are perfectly negatively correlated, i.e., $\rho_{12} = -1$, then $V(\alpha^*) = 0$. The expected rate of this portfolio is $(E(R_1) + E(R_2))/2$. So we are almost surely assured of the expected return portfolio.

V. Mean-variance optimization

As seen in the above section on diversification, it is possible with a judicious choice of portfolio allocation to diversify. But this type of risk reduction can be too severe and result in corresponding results that are smaller than a lesser diversified portfolio allocation. It is assumed that investors are willing to take on additional risk for a chance of getting higher returns. Harry Markowitz, an economist first formalized this trade-off and devised a systematic method for it. He even won the Nobel prize for this contribution.
The approach is given as: for a given level of expected returns, find the portfolio allocation with smallest risk. This can also be done by fixing the level of portfolio risk and then looking for the corresponding allocation that maximizes the expected return of this portfolio.
That is, we find a portfolio allocation $(\alpha_1, \alpha_2, \cdots, \alpha_N)$ so as to minimize the risk,

$$\frac{1}{2}\sigma^2\left(\sum_{i=1}^{M}\alpha_i R_i\right)$$

where, $\sum\limits_{i=1}^{M}\alpha_i = 1$, $E\left(\sum\limits_{i=1}^{M}\alpha_i\right) = \mu$

μ a given value which is the desired expected portfolio return. In practice α_i is range limited as $l_i \leq \alpha_i \leq u_i$ where l_i can be negative and u_i is generally non negative.

A. Stable probability models for finance

A stable distribution function for any two random variables X_1 and X_2 is defined as follows:

$$\gamma X + \delta = \alpha X_1 + \beta X_2$$

Here, it must be noted that the random variable X has the same distribution as X_1 and X_2, and γ and δ are constants.

The above concept can be applied to an analysis of business as follows:

Let N be the number of business days in a given month. Suppose $t_1, t_2, ..., t_N$ represent the times at which the stock market closes on the successive business days in that month. Let t_0 correspond to the last closing time in the preceding month. Then for $1 \leq n \leq N, S_n, r_n$, and ξ_n represent, respectively, the closing price, the return, and the log return for day (period) n.
We can recall that $\xi_n = \log(1+r_n) = \log(P_n) - \log(P_{n-1})$.
Let S_0 be the stock price at closing on the last business day of the preceding month.

We now have two ways of estimating the distribution of the monthly log return Ψ_N :

1) To estimate the distribution of the daily log using daily returns over 5 years, say, and then find ξ_n. The other is to estimate this distribution directly by using monthly returns, through closing prices at the end of each month for some Y years.

We observe that there are $Y \times 12 \times N$ daily returns over the five years. These observations help us decide on the distribution of the daily returns r_n, and thus we can also compute their logarithms ξ_n, and finally that of the monthly returns via their logarithm Psi_N, where Psi_N is given as follows.

$$\Psi_N = \sum_{i=1}^{N} \xi_i$$

2) The second approach is to observe $Y \times 12$ monthly returns over the Y years and infer the distribution of monthly returns (equivalent to the logarithm Ψ_N).

The difference between the two approaches can be noted as follows: with the first, we use actual data to infer the distribution of the daily returns. The distribution of the monthly returns is then obtained as a result of the addition of N random variables (the daily returns).

With the second approach, the distribution of the monthly returns is inferred directly from the data, not as a sum of the N random variables (daily returns).

Ideally, the latter distribution should be close to that which results from the sum of the daily returns. This property is known as *stability*.

VI. Black-Scholes Model

The *Black-Scholes Model* is a mathematical model used for understanding the dynamics of the financial market, including derivative investment instruments. Here, derivative is a contract that derives its value from the performance of an underlying asset.

The model consists of a partial differential equation known as the *Black-Scholes equation*, from which one can find the theoretical estimate of the price of European-style options (may be exercised only at the expiration date of the option, i.e. at a single pre-defined point in time) and shows that and option has a unique price when we already know the risk and expected return of that option.

A. Black-Scholes Equation

The Black-Scholes equation is:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \cdot \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

where $S(t)$ refers to the price of the underlying asset at time t, $V(S,t)$ is the price of an option as a function of the underlying asset S at time t. We can use this equation to perfectly offset any potential losses or gains that may be incurred by a companion investment by buying and selling the underlying and the bank asset in just the right way, and thus make the risk zero.

B. Black-Scholes formula

Some terms before proceeding,

- S: Current stock price
- K: Strike price
- r: Risk-free rate
- σ: Volatility
- T: Time of option expiration
- t: A time (in years)
- $\tau = T - t$: Time to maturity

This formula calculates the price of the European put and call options. It is obtained by applying the terminal and boundary conditions to the above equation as follows:

- The European call option, $C(0,t) = 0 \ \forall \ t$
- $C(S,t) \to S$ as $S \to \infty$
- $C(S,t) = \max\{S - K, 0\}$

Before looking at the formula, we define the standard normal cumulative distribution function $N(x)$ as

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dz$$

We define the following,

$$d_1 = \frac{1}{\sigma\sqrt{T-t}} \left[\ln \frac{S_t}{K} + \left(r + \frac{\sigma^2}{2}(T-t) \right) \right]$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

The value of a *call option* for an underlying stock which is non-dividend paying in terms of the defined parameters is stated below using the terms:

$$\boxed{C(S_t, t) = N(d_1)S_t - N(d_2) \cdot Ke^{-r(T-t)}}$$

Similarly, the value of a *put option* based on put-call parity with discount factor $e^{-r(T-t)}$ is

$$\boxed{\begin{aligned} P(S_t, t) &= Ke^{-r(T-t)} - S_t + C(S_t, t) \\ &= N(-d_2)Ke^{-r(T-t)} - N(-d_1)S_t \end{aligned}}$$

C. Code

```python
import numpy as np
from scipy.stats import norm

# Implementing Black-Scholes Model
def BlackScholes(S, K, r, sigma, t, option):
    # S: Current stock price
    # K: Strike price
    # r: Risk-free rate
    # sigma: Volatility
    # t: Time to maturity
    # option: 'call' or 'put'

    d1 = (np.log(S/K) + (r + (sigma**2)/2) * t) / (sigma * np.sqrt(t))
    d2 = d1 - sigma * np.sqrt(t)

    # Call option
    if option == 'call':
        return S * norm.cdf(d1) - K * np.exp(-r*t) * norm.cdf(d2)

    # Put option
    if option == 'put':
        return K * np.exp(-r*t) * norm.cdf(-d2) - S * norm.cdf(-d1)
```

Fig. 5. Black-Scholes Model Implementation

References

[1] K. L. Chung, and Farid AitSahlia, “Elementary Probability Theory: With Stochastic Processes and an Introduction to Mathematical Finance (Undergraduate Texts in Mathematics)”

[2] Black-Scholes Model - Investopedia

[3] Black-Scholes model - Wikipedia, the Free Encyclopedia

[4] Black-Scholes Formula and Python Implementation

VII. Contributions

1) **Aditya Nair**
   Asset Returns, Asset Risk, Black-Scholes Model
2) **Agrim Rawat**
   Introduction, Black-Scholes Model, Code
3) **Sankalp S. Bhat**
   Portfolios, Asset Returns, Asset Risk
4) **Tejah S S**
   Introduction, Portfolios, Mean Variance Optimization