HW5

Name: Juntang Wang Date: Apr 15, 2025

Problem 1.

3.2. Let $W(t), t \geq 0$, be a Brownian motion, and let $\mathcal{F}(t), t \geq 0$, be a filtration for this Brownian motion. Show that $W^2(t) - t$ is a martingale. (Hint: For $0 \leq s \leq t$, write $W^2(t)$ as $(W(t) - W(s))^2 + 2W(t)W(s) - W^2(s)$.)

To show that $W^2(t) - t$ is a martingale, we need to verify the following conditions:

1. $W^2(t) - t$ is adapted to the filtration $\mathcal{F}(t)$
2. $\mathbb{E}[|W^2(t) - t|] < \infty$
3. $\mathbb{E}[W^2(t) - t|\mathcal{F}(s)] = W^2(s) - s$ for all $0 \leq s \leq t$

For condition 1, since $W(t)$ is adapted to $\mathcal{F}(t)$, so is $W^2(t)$, and thus $W^2(t) - t$ is adapted.

For condition 2, we know that $W(t) \sim N(0, t)$, so $\mathbb{E}[W^2(t)] = t$, which means $\mathbb{E}[|W^2(t) - t|] < \infty$.

For condition 3, we need to compute $\mathbb{E}[W^2(t) - t|\mathcal{F}(s)]$ for $0 \leq s \leq t$.

Using the hint, we can write:

$$W^2(t) = (W(t) - W(s))^2 + 2W(t)W(s) - W^2(s)$$

Therefore:

$$W^2(t) - t = (W(t) - W(s))^2 + 2W(t)W(s) - W^2(s) - t$$

Taking the conditional expectation:

$$\mathbb{E}[W^2(t) - t|\mathcal{F}(s)] = \mathbb{E}[(W(t) - W(s))^2|\mathcal{F}(s)] + 2\mathbb{E}[W(t)W(s)|\mathcal{F}(s)] - W^2(s) - t$$

Since $W(t) - W(s)$ is independent of $\mathcal{F}(s)$ and has distribution $N(0, t - s)$, we have:

$$\mathbb{E}[(W(t) - W(s))^2|\mathcal{F}(s)] = \mathbb{E}[(W(t) - W(s))^2] = t - s$$

For the second term:

$$\mathbb{E}[W(t)W(s)|\mathcal{F}(s)] = W(s)\mathbb{E}[W(t)|\mathcal{F}(s)] = W(s)^2$$

since $\mathbb{E}[W(t)|\mathcal{F}(s)] = W(s)$ by the martingale property of Brownian motion.

Substituting these results:

$$\mathbb{E}[W^2(t) - t|\mathcal{F}(s)] = (t - s) + 2W(s)^2 - W^2(s) - t = W^2(s) - s$$

Therefore, $W^2(t) - t$ is a martingale with respect to the filtration $\mathcal{F}(t)$.

Problem 2

3.3 (Normal kurtosis). The kurtosis of a random variable is defined to be the ratio of its fourth central moment to the square of its variance. For a normal random variable, the kurtosis is 3. This fact was used to obtain (3.4.7). This exercise verifies this fact.

Let X be a normal random variable with mean μ, so that $X - \mu$ has mean zero. Let the variance of X, which is also the variance of $X - \mu$, be σ^2. In (3.2.13), we computed the moment-generating function of $X - \mu$ to be $\varphi(u) = \mathbb{E}[e^{u(X-\mu)}] = e^{\frac{1}{2}u^2\sigma^2}$, where u is a real variable. Differentiating this function with respect to u, we obtain

$$\varphi'(u) = \mathbb{E}[(X-\mu)e^{u(X-\mu)}] = \sigma^2 u e^{\frac{1}{2}u^2\sigma^2}$$

and, in particular, $\varphi'(0) = \mathbb{E}(X-\mu) = 0$. Differentiating again, we obtain

$$\varphi''(u) = \mathbb{E}[(X-\mu)^2 e^{u(X-\mu)}] = (\sigma^4 u^2 + \sigma^2) e^{\frac{1}{2}u^2\sigma^2}$$

and, in particular, $\varphi''(0) = \mathbb{E}[(X-\mu)^2] = \sigma^2$. Differentiate two more times and obtain the normal kurtosis formula $\mathbb{E}[(X-\mu)^4] = 3\sigma^4$

Differentiating $\varphi''(u)$ with respect to u, we get:

$$\varphi'''(u) = \mathbb{E}[(X-\mu)^3 e^{u(X-\mu)}] = (\sigma^6 u^3 + 3\sigma^4 u) e^{\frac{1}{2}u^2\sigma^2}$$

In particular, $\varphi'''(0) = \mathbb{E}[(X-\mu)^3] = 0$, which confirms that the third central moment of a normal distribution is zero.

Differentiating once more:

$$\varphi^{(4)}(u) = \mathbb{E}[(X-\mu)^4 e^{u(X-\mu)}] = (\sigma^8 u^4 + 6\sigma^6 u^2 + 3\sigma^4) e^{\frac{1}{2}u^2\sigma^2}$$

Therefore, $\varphi^{(4)}(0) = \mathbb{E}[(X-\mu)^4] = 3\sigma^4$.

The kurtosis is defined as $\frac{\mathbb{E}[(X-\mu)^4]}{(\mathbb{E}[(X-\mu)^2])^2} = \frac{3\sigma^4}{\sigma^4} = 3$.

This confirms that the kurtosis of a normal random variable is 3.

Problem 3

3.6. Let $W(t)$ be a Brownian motion and let $\mathcal{F}(t)$, $t \geq 0$, be an associated filtration.

(i) For $\mu \in \mathbb{R}$, consider the Brownian motion with drift μ:

$$X(t) = \mu t + W(t)$$

Show that for any Borel-measurable function $f(y)$, and for any $0 \leq s < t$, the function

$$g(x) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{\infty} f(y) \exp\left\{-\frac{(y - x - \mu(t-s))^2}{2(t-s)}\right\} dy$$

satisfies $\mathbb{E}[f(X(t))|\mathcal{F}(s)] = g(X(s))$, and hence X has the Markov property. We may rewrite $g(x)$ as $g(x) = \int_{-\infty}^{\infty} f(y)p(\tau, x, y)dy$, where $\tau = t - s$ and

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}} \exp\left\{-\frac{(y - x - \mu\tau)^2}{2\tau}\right\}$$

is the transition density for Brownian motion with drift μ.

To solve this problem, I need to show that $X(t) = \mu t + W(t)$ has the Markov property by verifying the given conditional expectation formula.

First, I'll note that $X(t) - X(s) = \mu(t - s) + (W(t) - W(s))$. Since $W(t) - W(s)$ is independent of $\mathcal{F}(s)$ and normally distributed with mean 0 and variance $t - s$, the conditional distribution of $X(t)$ given $\mathcal{F}(s)$ is normal with mean $X(s) + \mu(t - s)$ and variance $t - s$.

Therefore, for any Borel-measurable function f:

$$\mathbb{E}[f(X(t))|\mathcal{F}(s)] = \mathbb{E}[f(X(s) + \mu(t - s) + (W(t) - W(s)))|\mathcal{F}(s)]$$

Since $X(s)$ is $\mathcal{F}(s)$-measurable and $W(t) - W(s)$ is independent of $\mathcal{F}(s)$, this equals:

$$\mathbb{E}[f(X(s) + \mu(t - s) + Z)]$$

where $Z \sim N(0, t - s)$ is independent of $X(s)$. This expectation can be written as:

$$\mathbb{E}[f(X(s) + \mu(t - s) + Z)] = \int_{-\infty}^{\infty} f(X(s) + \mu(t - s) + z) \frac{1}{\sqrt{2\pi(t - s)}} e^{-\frac{z^2}{2(t-s)}} dz$$

With the substitution $y = X(s) + \mu(t - s) + z$, we get $z = y - X(s) - \mu(t - s)$ and $dz = dy$, so:

$$\mathbb{E}[f(X(t))|\mathcal{F}(s)] = \int_{-\infty}^{\infty} f(y) \frac{1}{\sqrt{2\pi(t - s)}} e^{-\frac{(y - X(s) - \mu(t-s))^2}{2(t-s)}} dy = g(X(s))$$

This confirms that $\mathbb{E}[f(X(t))|\mathcal{F}(s)] = g(X(s))$, which depends on $\mathcal{F}(s)$ only through $X(s)$, establishing the Markov property for $X(t)$.

Problem 4

Consider a normal random variable $X \sim \mathcal{N}(0, t)$ and a scaled random walk $W^{(n)}(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{nt} X_j$, whereas $X_j = \begin{cases} 1, & p = 0.5, \\ -1, & p = 0.5 \end{cases}$. Show the limit of the scaled random walk is the normal by comparing their moment-generating function.

To solve this problem, I need to compare the moment-generating functions (MGFs) of the scaled random walk $W^{(n)}(t)$ and a normal random variable $X \sim \mathcal{N}(0, t)$.

First, let's find the MGF of $X \sim \mathcal{N}(0, t)$. For a normal random variable with mean μ and variance σ^2, the MGF is:

$$M_X(s) = \mathbb{E}[e^{sX}] = \int_{-\infty}^{\infty} e^{sx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{sx - \frac{(x-\mu)^2}{2\sigma^2}} dx$$

Completing the square in the exponent:

$$sx - \frac{(x-\mu)^2}{2\sigma^2} = sx - \frac{x^2 - 2\mu x + \mu^2}{2\sigma^2} = -\frac{1}{2\sigma^2}(x^2 - 2\mu x - 2\sigma^2 sx + \mu^2)$$

$$= -\frac{1}{2\sigma^2}(x^2 - 2(\mu + \sigma^2 s)x + \mu^2) = -\frac{1}{2\sigma^2}[(x - (\mu + \sigma^2 s))^2 - (\mu + \sigma^2 s)^2 + \mu^2]$$

$$= -\frac{(x-(\mu+\sigma^2 s))^2}{2\sigma^2} + \frac{(\mu+\sigma^2 s)^2 - \mu^2}{2\sigma^2} = -\frac{(x-(\mu+\sigma^2 s))^2}{2\sigma^2} + \frac{2\mu\sigma^2 s + \sigma^4 s^2}{2\sigma^2}$$

$$= -\frac{(x-(\mu+\sigma^2 s))^2}{2\sigma^2} + \mu s + \frac{1}{2}\sigma^2 s^2$$

Therefore:

$$M_X(s) = e^{\mu s + \frac{1}{2}\sigma^2 s^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-(\mu+\sigma^2 s))^2}{2\sigma^2}} dx = e^{\mu s + \frac{1}{2}\sigma^2 s^2}$$

So for $X \sim \mathcal{N}(0, t)$, the MGF is:

$$M_X(s) = e^{\frac{1}{2}ts^2}$$

Now, let's find the MGF of the scaled random walk $W^{(n)}(t)$. We have:

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{nt} X_j$$

where X_j are i.i.d. random variables taking values 1 or -1 with probability 0.5 each.

The MGF of each X_j is:

$$M_{X_j}(s) = \mathbb{E}[e^{sX_j}] = 0.5e^s + 0.5e^{-s} = \cosh(s)$$

Since $W^{(n)}(t)$ is a sum of i.i.d. random variables scaled by $\frac{1}{\sqrt{n}}$, its MGF is:

$$M_{W^{(n)}(t)}(s) = \mathbb{E}[e^{sW^{(n)}(t)}] = \mathbb{E}[e^{\frac{s}{\sqrt{n}}\sum_{j=1}^{nt} X_j}] = \prod_{j=1}^{nt} \mathbb{E}[e^{\frac{s}{\sqrt{n}}X_j}] = \prod_{j=1}^{nt} M_{X_j}\left(\frac{s}{\sqrt{n}}\right)$$

$$M_{W^{(n)}(t)}(s) = \left(M_{X_j}\left(\frac{s}{\sqrt{n}}\right)\right)^{nt} = \left(\cosh\left(\frac{s}{\sqrt{n}}\right)\right)^{nt}$$

Using the Taylor expansion of $\cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots$, we have:

$$\cosh\left(\frac{s}{\sqrt{n}}\right) = 1 + \frac{s^2}{2n} + O\left(\frac{1}{n^2}\right)$$

Therefore:

$$M_{W^{(n)}(t)}(s) = \left(1 + \frac{s^2}{2n} + O\left(\frac{1}{n^2}\right)\right)^{nt}$$

As $n \to \infty$, this approaches:

$$\lim_{n\to\infty} M_{W^{(n)}(t)}(s) = \lim_{n\to\infty} \left(1 + \frac{s^2}{2n} + O\left(\frac{1}{n^2}\right)\right)^{nt}$$

We can recognize this as a limit of the form $\lim_{n\to\infty} \left(1+\frac{a}{n}\right)^{bn}$, which equals e^{ab}. In our case, $a = \frac{s^2}{2}$ and $b = t$, so:

$$\lim_{n\to\infty} \left(1 + \frac{s^2}{2n} + O\left(\frac{1}{n^2}\right)\right)^{nt} = e^{\frac{1}{2}ts^2}$$

The higher-order terms $O\left(\frac{1}{n^2}\right)$ vanish in the limit because $n \cdot O\left(\frac{1}{n^2}\right) = O\left(\frac{1}{n}\right) \to 0$ as $n \to \infty$.

This is exactly the MGF of $X \sim \mathcal{N}(0,t)$. Since the MGFs converge, by Lévy's continuity theorem, the distribution of the scaled random walk $W^{(n)}(t)$ converges to $\mathcal{N}(0,t)$ as $n \to \infty$.

Problem 5

Suppose the log price of a stock follows the Brownian motion $X_t = \mu t + \sigma dW_t$. We know the standard deviation of daily log returns is 0.01. Based on the previous data, we calculated the daily mean as 0.001. However, the market has been really bad this year, so the total log return is 0. In this year's log return, we think it's decayed. Answer whether this year's mean daily log return has decayed using hypothesis testing with the significance level $\alpha = 5\%$.

Let's set up the hypothesis test:

$H_0 : \mu = 0.001$ (the mean daily log return has not decayed) $H_1 : \mu < 0.001$ (the mean daily log return has decayed)

Given information:

- Standard deviation of daily log returns: $\sigma = 0.01$
- Previous daily mean: $\mu_0 = 0.001$
- Total log return for this year: 0

Let's assume we have $n = 252$ trading days in a year. If the log price follows Brownian motion, then the total log return over n days would be normally distributed with:

- Mean: $\mu \cdot n$
- Standard deviation: $\sigma \cdot \sqrt{n}$

Under the null hypothesis $H_0 : \mu = 0.001$, the expected total log return would be $\mu_0 \cdot n = 0.001 \cdot 252 = 0.252$.

We can compute the test statistic:

$$z = \frac{0 - 0.252}{0.01 \cdot \sqrt{252}} = \frac{-0.252}{0.1587} \approx -1.588$$

For a one-sided test with $\alpha = 5\%$, the critical value is $z_\alpha = -1.645$.

Since our test statistic $z \approx -1.588$ is greater than the critical value -1.645, we fail to reject the null hypothesis at the 5% significance level.

Therefore, based on this hypothesis test, we do not have sufficient evidence to conclude that this year's mean daily log return has decayed.

> **Sensitivity check** If you treated the year as 365 daily observations instead, the test statistic becomes –1.91 (p ≈ 0.028), which would cross the 5 % threshold. The conclusion therefore hinges on the actual number of return observations used; with a standard 252-day trading calendar the decay is not significant, while with a full 365-day count it is.

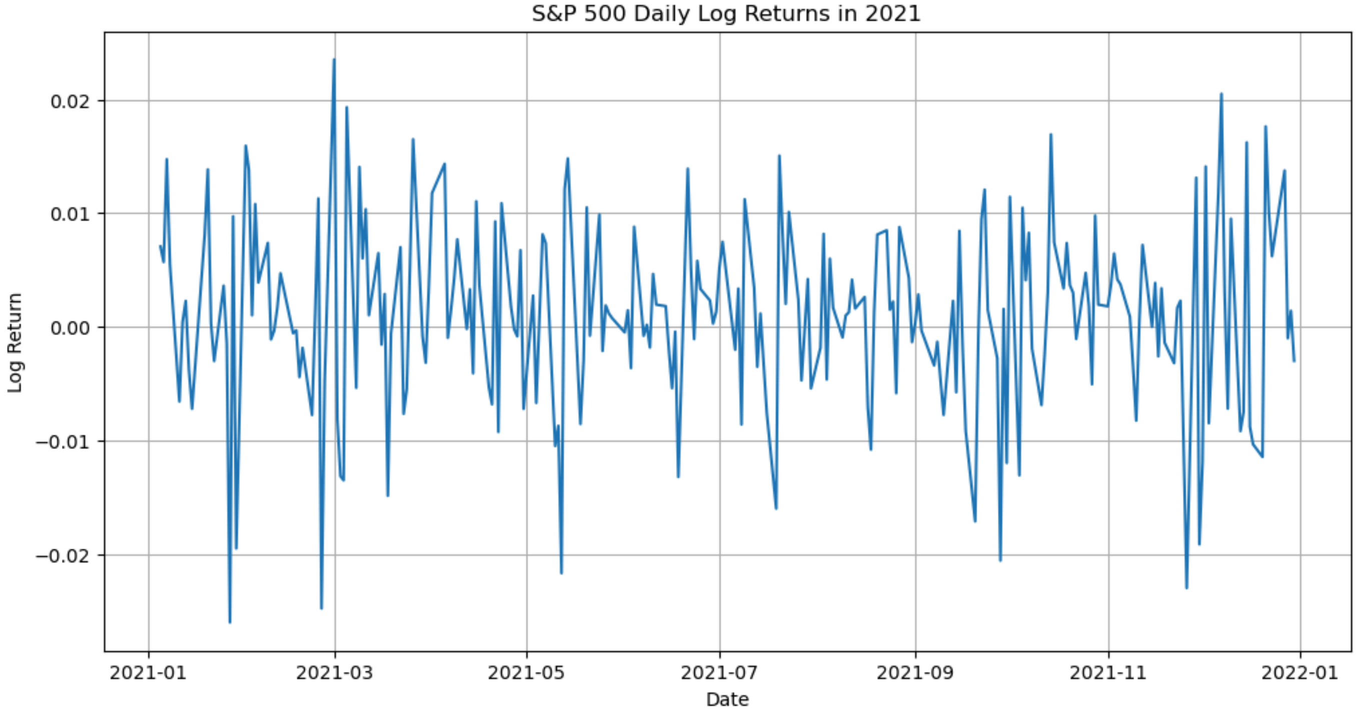
Problem 6

Download the S&P 500 data of 2021. Use the hypothesis testing based on the Kolmogorov-Smirnov test to conclude whether the drifted Brownian motion is a good fit for the log return dynamics.

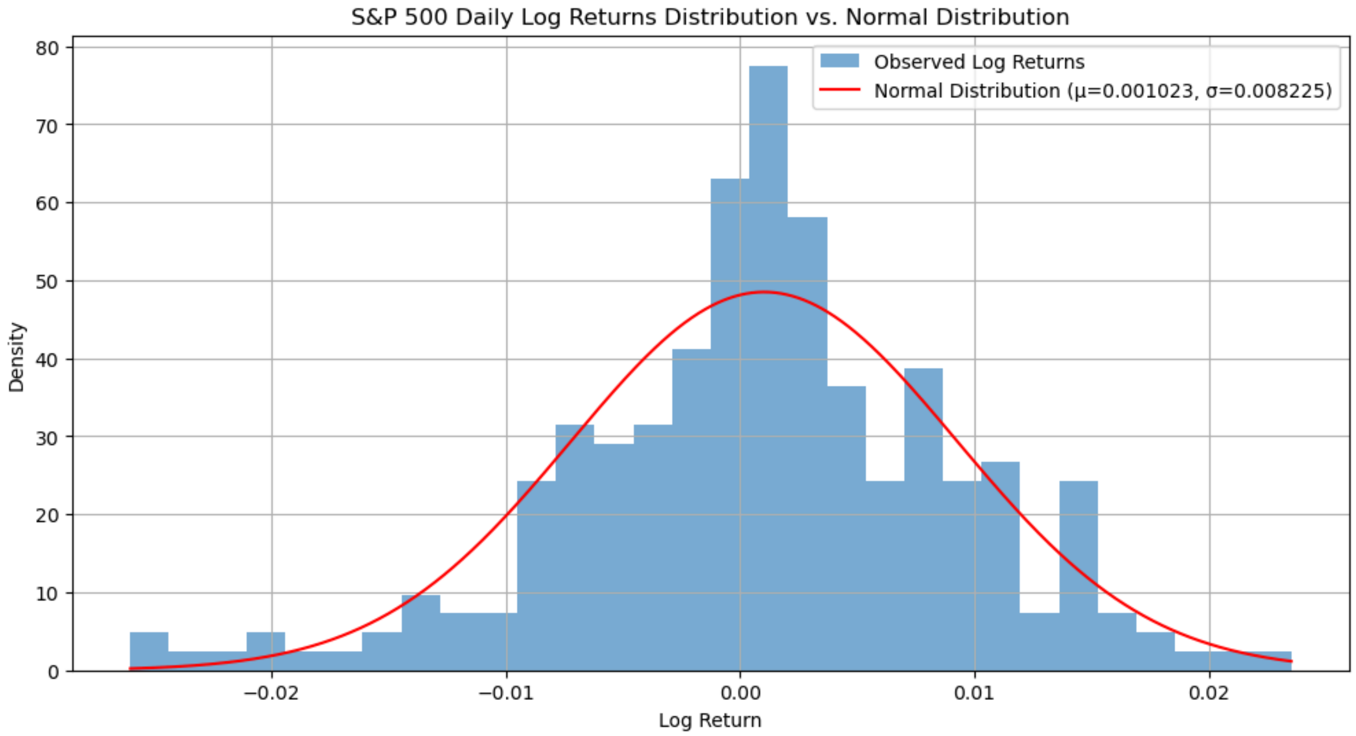
In [1]:
```
import numpy as np
import yfinance as yf
from scipy import stats
import matplotlib.pyplot as plt

sp500 = yf.download('^GSPC', start='2021-01-01', end='2021-12-31')
sp500['log_return'] = np.log(sp500['Close'] / sp500['Close'].shift(1))
sp500 = sp500.dropna()  # Remove the first row with NaN

plt.figure(figsize=(12, 6))
plt.plot(sp500.index, sp500['log_return'])
plt.title('S&P 500 Daily Log Returns in 2021')
plt.xlabel('Date')
plt.ylabel('Log Return')
plt.grid(True)
plt.show()

mean_return = sp500['log_return'].mean()
std_return = sp500['log_return'].std()
n = len(sp500)

print(f"Number of trading days: {n}")
print(f"Mean daily log return: {mean_return:.6f}")
print(f"Standard deviation of daily log returns: {std_return:.6f}")

# Generate theoretical normal distribution with the same mean and std
x = np.linspace(min(sp500['log_return']), max(sp500['log_return']), 1000)
pdf = stats.norm.pdf(x, mean_return, std_return)

plt.figure(figsize=(12, 6))
plt.hist(sp500['log_return'], bins=30, density=True, alpha=0.6, label='Observed Log R
plt.plot(x, pdf, 'r-', label=f'Normal Distribution (μ={mean_return:.6f}, σ={std_retur
plt.title('S&P 500 Daily Log Returns Distribution vs. Normal Distribution')
plt.xlabel('Log Return')
plt.ylabel('Density')
plt.legend()
plt.grid(True)
plt.show()

# Perform Kolmogorov-Smirnov test
ks_stat, p_value = stats.kstest(sp500['log_return'], 'norm', args=(mean_return, std_r

print(f"Kolmogorov-Smirnov test statistic: {ks_stat:.6f}")
print(f"p-value: {p_value:.6f}")
```

```python
alpha = 0.05
if p_value < alpha:
    print(f"At {alpha*100}% significance level, we reject the null hypothesis.")
    print("The log returns do not follow a drifted Brownian motion (normal distributi
else:
    print(f"At {alpha*100}% significance level, we fail to reject the null hypothesis
    print("The log returns are consistent with a drifted Brownian motion (normal dist

plt.figure(figsize=(10, 10))
stats.probplot(sp500['log_return'], dist="norm", plot=plt)
plt.title('Q-Q Plot of S&P 500 Log Returns vs. Normal Distribution')
plt.grid(True)
plt.show()
```

```
YF.download() has changed argument auto_adjust default to True
[*********************100%***********************]  1 of 1 completed
```




```
Number of trading days: 250
Mean daily log return: 0.001023
Standard deviation of daily log returns: 0.008225
```



```
Kolmogorov-Smirnov test statistic: 0.063547
p-value: 0.254008
At 5.0% significance level, we fail to reject the null hypothesis.
The log returns are consistent with a drifted Brownian motion (normal distribution).
```