

# Financial Market Distributions

Reference: Bodie et al, Ch 5

Econ 457

Week 2-a

# Outline

---

1. The Normal Distribution
2. Distributions of Financial Market Returns
3. Annualizing Weekly Returns
4. Forecasting
5. Problems with Using the Normal Distribution
6. Risk management enhancements

# 1. The Normal Distribution

---

Some basic facts

Notation:

$$X \sim N(\mu, \sigma^2)$$

Where  $\mu$  is the mean,  $\sigma^2$  is the variance, and  $\sigma$  is the standard deviation.

The distribution can be characterized by either the variance or the standard deviation. In finance it is common to use the standard deviation, because it is in the same units as the returns, whereas the variance is squared.

# 1. The Normal Distribution

Some basic facts

Empirical mean:

$$\bar{x} = \frac{1}{T} \sum_{t=1}^T x_t$$

Empirical variance:

$$s^2 = \frac{1}{T-1} \sum_{t=1}^T (x_t - \bar{x})^2$$

Empirical standard deviation:

$$s = \sqrt{s^2} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (x_t - \bar{x})^2}$$

# 1. The Normal Distribution

---

Some basic facts, cont'd

Adding a constant  $C$  to  $X$ :

$$X + C \sim N(\mu_x + C, \sigma_x^2)$$

Multiplying  $X$  by a constant  $C$ :

$$C \cdot X \sim N(C\mu_x, C^2\sigma_x^2)$$

(notice the  $C^2$  in the variance term)

# 1. The Normal Distribution

Some basic facts, cont'd

Adding two normally distributed variables  $X$  and  $Y$ :

**Case 1:**  $X$  and  $Y$  are *independent*

$$X + Y \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$$

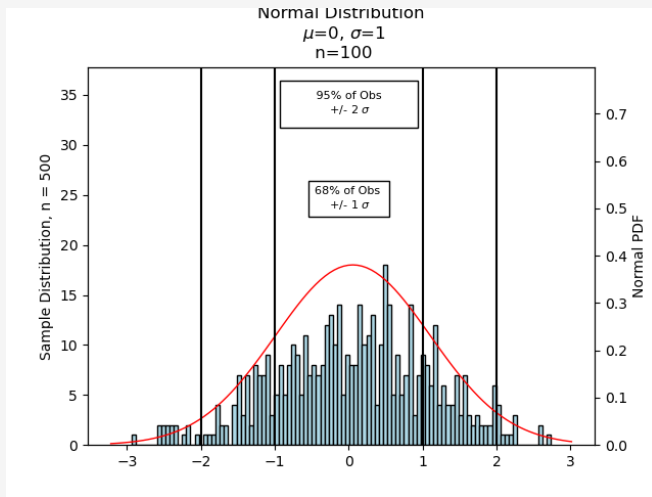
**Case 2:**  $X$  and  $Y$  are *correlated*

$$X + Y \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2 + 2\sigma_x\sigma_y\rho_{x,y})$$

where  $\rho_{x,y} = \frac{\text{Cov}(X,Y)}{\sigma_x\sigma_y}$  is the correlation coefficient.

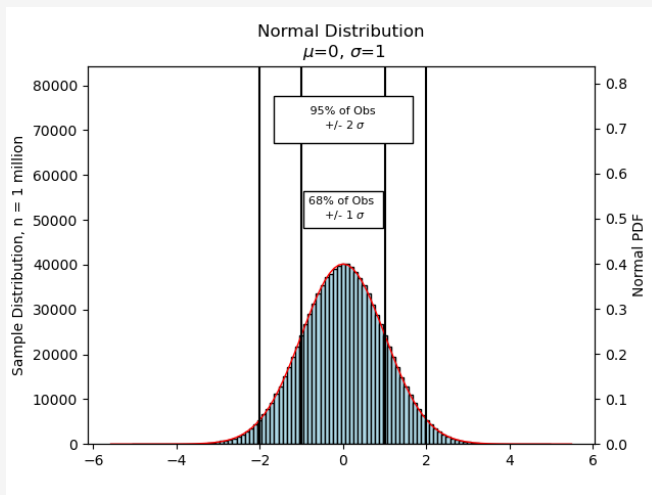
# 1. The Normal Distribution

A Sample,  $n = 100$



# 1. The Normal Distribution

Some basic facts, cont'd



# 1. The Normal Distribution

Some basic facts, cont'd

**The Central Limit Theorem** The distribution of the sample mean converges to a normal distribution, regardless of the distribution of the underlying variable. Let  $\bar{X} = \frac{1}{T} \sum_i^T x_i$  be the sample mean, then

$$\bar{X} \xrightarrow{d} N\left(\mu, \frac{\sigma^2}{T}\right) \text{ as } T \rightarrow \infty$$

The only requirement is that the underlying variables (usually) need to be independent and identically distributed.

# 1. The Normal Distribution

---

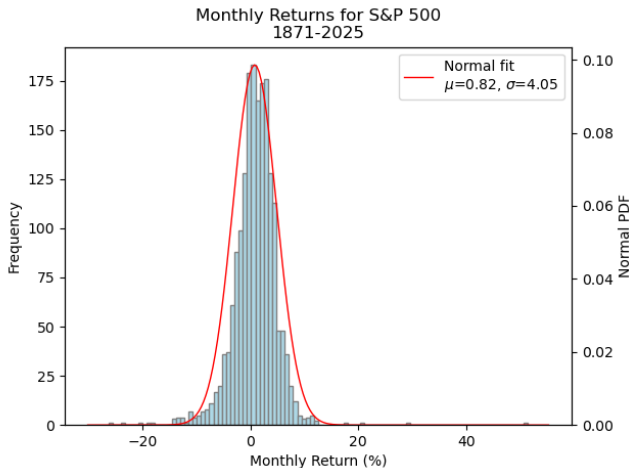
Some basic facts, cont'd

## **Galton Board**

<https://youtu.be/EvHiee7gs9Y>

## 2. Distributions of Financial Market Returns

S&P 500



Data source: Robert Shiller, [www.shillerdata.com](http://www.shillerdata.com)

## 2. Distributions of Financial Market Returns

---

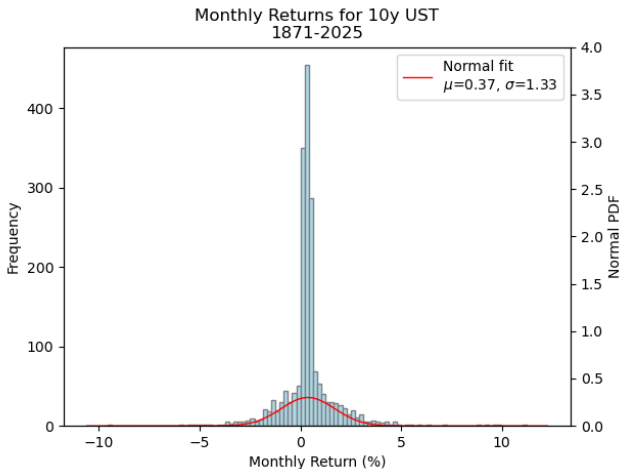
### S&P 500 Total Returns

Table: S&P 500 Total Returns by Decade, (% Annualized)

Decade	Return	Standard Deviation
1870s	8.11	10.86
1880s	6.29	9.62
1890s	6.24	11.84
1900s	10.62	12.70
1910s	4.99	10.51
1920s	15.46	15.08
1930s	4.39	30.63
1940s	9.47	13.23
1950s	18.24	10.08
1960s	8.05	10.31
1970s	6.56	13.23
1980s	16.88	12.98
1990s	17.19	10.45
2000s	0.38	14.68
2010s	13.02	9.55
2020s	13.72	14.15

## 2. Distributions of Financial Market Returns

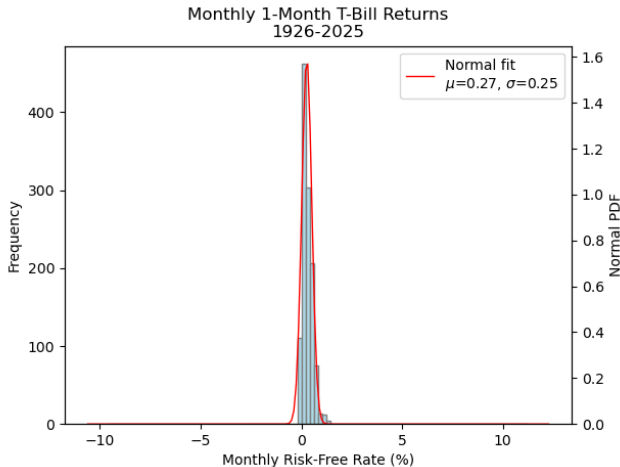
US Treasury Bonds, 10-year US Treasury



Data source: Robert Shiller, [www.shillerdata.com](http://www.shillerdata.com)

## 2. Distributions of Financial Market Returns

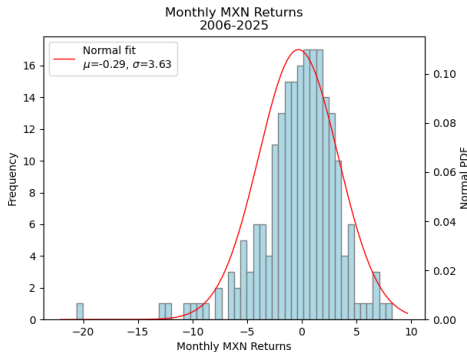
US T-Bills, 1-month



Data source: Ken French

## 2. Distributions of Financial Market Returns

MXN



Data source: Interactive Brokers (USDMXN Cash) Only the price return, doesn't include carry 5 worst months: Mar 2020, Sep 2011, Oct 2008, May 2012, Nov 2008

### 3. Annualizing Weekly Returns

---

If you are given a data at a higher frequency it is often convenient to convert the mean and standard deviation to annual equivalents.

In order to do so, imagine that each year is made up of *independent* sub-periods. Then we can use the following property of the normal distribution:

$$\begin{aligned}x_1 + \dots + x_T &\sim N(\mu_x + \dots + \mu_x, \sigma_x^2 + \dots + \sigma_x^2) \\ &\sim N(T\mu_x, T\sigma_x^2)\end{aligned}$$

Where  $x_1$  through  $x_T$  are independent sub-periods, and  $T$  is the number of sub-periods that make up the longer period. (i.e. 52 weeks in a year, so  $T=52$ ) We assume the distributions of each sub-period are identical and equal to  $N(\mu_x, \sigma_x^2)$ .

### 3. Annualizing Weekly Returns

---

To compute annual returns from a series of weekly returns (52 weeks in a year):

1. Compute the mean and standard deviation of the weekly returns
2. Multiply the estimated mean by 52
  - An alternative would be to account for compounding and use  $(1 + \mu_x)^{52} - 1$ .
3. Multiply the standard deviation by  $\sqrt{52}$ 
  - Be careful here, this is because the variance increases by  $T$ , so standard deviation increases by  $\sqrt{T}$

Easy to do something similar for daily returns (multiply standard deviation by  $\sqrt{250}$ , number of trading days in a year), or monthly returns (multiply standard deviation by  $\sqrt{12}$ ), or quarterly returns (multiply standard deviation by 2).

### 3. Annualizing Weekly Returns

---

Table: Daily, Monthly, Annual S&P Returns

	Mean	Standard Deviation
Daily	0.04	1.19
Daily, annualized	9.48	18.87
Monthly	0.75	4.25
Monthly, annualized	8.99	14.73
Annual	9.35	16.31

---

Data from Interactive Brokers, since 2004

## 4. Forecasting

---

Note that the discussion up to now has been about measuring historical returns. In contrast, the questions we usually care about concern expected future returns.

In general, you should be very careful about using historical returns and distributions to forecast future returns and distributions.

But forecasting is an inescapable requirement for many things in financial markets, so you have to come up with something. Much of the next few weeks will be focused on how to think about *expected* returns and prices for different securities.

## 4. Forecasting

---

If you think history is a good guide...

If you can convince yourself that history is a good guide for future returns (or if you can find a historical period that is a good guide), then can use the empirical mean and standard deviation:

$$\mathbb{E}(r) = \frac{1}{T} \sum_i^T r_i$$

$$\text{Var}(r) = \frac{1}{T} \sum_i^T (r_i - \mathbb{E}[r])^2$$

where  $r_i$  are *historical* returns measured over  $T$  periods.

## 4. Forecasting

---

If you can guess the probability of scenarios...

If we start with a set of estimates for returns in different scenarios and the probabilities of those scenarios, then we can compute the expected value and expected variance as follows:

$$\mathbb{E}(r) = \sum_i p_i \cdot r_i$$

$$\text{Var}(r) = \sum_i p_i \cdot (r_i - \mathbb{E}[r])^2$$

where  $p_i$  is the probability and  $r_i$  is the return in scenario  $i$ .

Note these are analogous to the empirical estimates of the mean and the variance discussed earlier.

## 5. Problems with Using the Normal Distribution

---

While it may be convenient to use the normal distribution for forecasts of future returns, doing so raises a lot of problems:

1. Observations may not be independent (Central Limit Theorem doesn't hold)
2. Distributions of financial returns often have 'fat tails'
3. Historical returns have no measure of value
4. The variance of returns seems to change over time

What to do about this?

## 5. Problems with Using the Normal Distribution

---

Observations are not independent

### **Galton Board, v2**

<https://youtu.be/3m4bxse2JEQ?si=SIhl4vECWSKwkQV->

## 5. Problems with Using the Normal Distribution

### Fat Tails

**Table:** S&P Highest and Lowest Monthly Returns Since 1950

Date	Return %	Probability %
2009-04	12.32	0.05
1982-09	12.10	0.07
1991-02	11.61	0.11
1998-11	10.97	0.19
1975-02	10.81	0.23
1962-06	-11.41	0.02
1987-10	-11.85	0.01
1987-11	-12.30	0.01
2020-03	-18.92	0.00
2008-10	-20.19	0.00

**Table:** 10y UST Highest and Lowest Monthly Returns Since 1950

Date	Return %	Probability %
1981-11	11.23	0.00
1982-10	9.86	0.00
2008-12	9.74	0.00
1980-05	9.15	0.00
1980-04	8.87	0.00
2022-09	-5.18	0.16
2003-07	-5.30	0.13
2022-04	-5.42	0.10
1979-10	-5.85	0.05
1980-02	-9.56	0.00

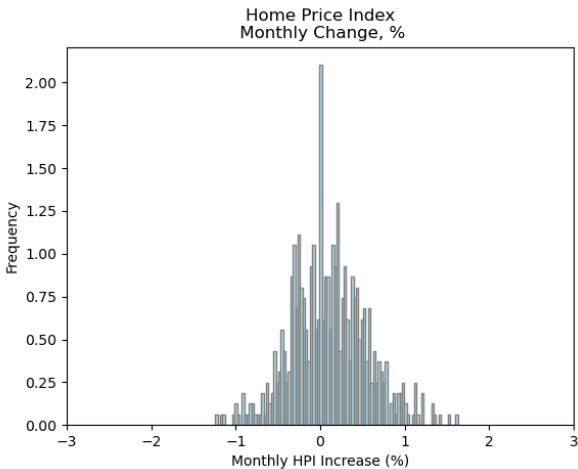
Probabilities are based on estimated normal distribution.  $n = 900$  months, S&P:  $\mu = 0.96$  and  $\sigma = 3.4$ , 10y UST:  $\mu = 0.42$  and  $\sigma = 1.9$

## 5. Problems with Using the Normal Distribution

---

Historical returns have no measure of value

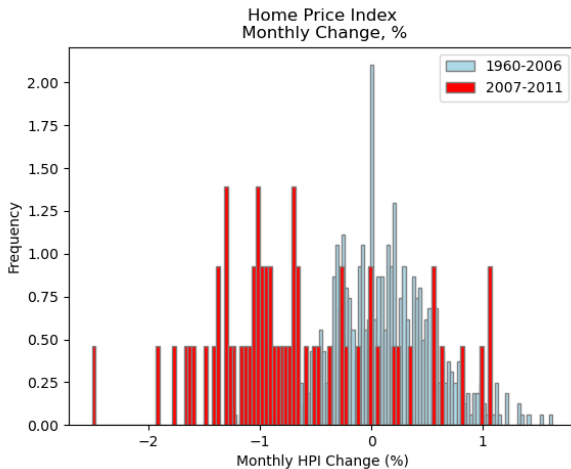
### Home prices, 1960-2006



## 5. Problems with Using the Normal Distribution

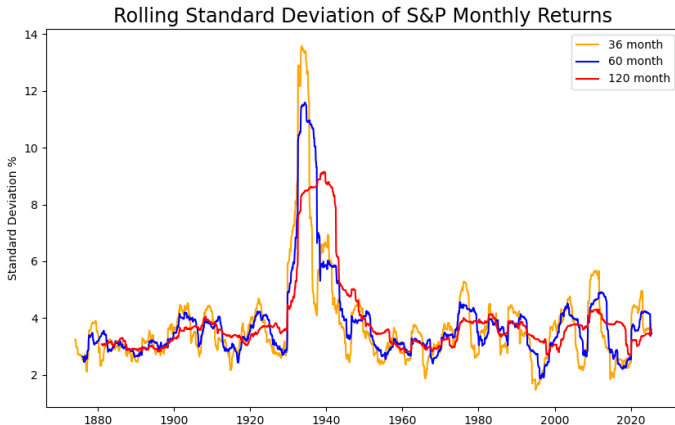
Historical returns have no measure of value

### Home prices, 1960-2025



## 5. Problems with Using the Normal Distribution

Variance changes over time



## 6. Risk management enhancements

---

What to do?

How financial market participants address these problems:

1. Use a distribution with fatter tails
2. Value at Risk (VaR)
3. Scenario Analysis / Stress Tests

## 6. Risk management enhancements

---

### Value at Risk (VaR) Definition

**Value at Risk (VaR)** measures the maximum expected loss over a specific time period at a given confidence level.

**Definition:**

- VaR answers: "What is the worst loss we can expect with X% confidence over Y days?"
- Example: 1-day 95% VaR = \$1 million means there's a 5% chance of losing more than \$1M tomorrow

## 6. Risk management enhancements

---

### Value at Risk (VaR) - Calculation

#### Calculation Methods:

1. **Parametric (Normal) VaR:** Assumes returns follow normal distribution

$$\text{VaR} = \mu - z_{\alpha} \times \sigma \times \sqrt{t}$$

where  $z_{\alpha}$  is the critical value (e.g., 1.65 for 95% confidence)

2. **Historical VaR:** Uses actual historical return distribution
3. **Monte Carlo VaR:** Simulates thousands of possible outcomes

**Limitations:** Still need to make an assumption about the tails, past patterns? what else?

## 6. Risk management enhancements

---

### Stress Tests

**Stress Testing:** Evaluates portfolio performance under extreme scenarios

**Approach:**

- Apply historical crisis scenarios (2008, 1987, COVID-19)
- Test hypothetical extreme events (interest rate shocks, market crashes)
- Examine correlation breakdown during crises

**Example:** "What would happen to our portfolio if we had another 2008-style crisis?"

Useful at the portfolio level, but need to specify correlations in addition to returns of various securities.