Uncertainty, Default, and Risk

(Welch, Chapter 06-A)

Ivo Welch

Mon Jun 20 14:48:30 2022

1: Maintained Assumptions

Perfect Markets

1. No differences in opinion.
2. No taxes.
3. No transaction costs.
4. No big sellers/buyers—infinitely many clones that can buy or sell.

BUT NO LONGER Certainty

2: Sadistics

- We now need to predict (describe) the future.
- For this, we need statistics.

3: Legal Disclaimer

This chapter may be illegal in some states. Attend the rest of this course at your own risk.

For example:

Persons pretending to forecast the future shall be considered disorderly under subdivision 3, section 901 of the criminal code and liable to a fine of $ 250 and/or six months in prison.

(Section 889, New York State Code of Criminal Procedure.)

4: Statistics

- Covered more fully in your sadistics course.
- This brisk overview is just the minimum needed.

5: Random Variables (RVs)

- A **Random Variable (RV)** (often with tilde over the variable) is
  - not an ordinary algebra variable,
    - that has some specific unknown value,
  - but best thought-of as a histogram with a name,
    - a "cloud" outcome from a randomizing device (e.g., a coin or a card draw), whose histogram defines it.

6: Inference

- In many applications, statisticians *assume* you understand the system and know the histogram, but not the draw (the outcome).
  - **Great for**: coin, die, roulette.
  - **Not sure for**: financial markets (where we do not understand the underlying physics).
- On occasion, I will warn you.

7: Die Throw RV

- Our main random variable example will be the payoff you get after *a* die is thrown:
  - 1 = $-\$6$,
  - 2 = $+\$36$,
  - 3 = $-\$12$,
  - 4 to 6 = $+\$150$.

Let's call this random variable D (ie $\tilde{D}$).

- Draw its histogram.

8: Draw the Histogram of D

9: The Expected Value

- $E(\cdot)$ is common notation for *Expected Value* (EV).
  - Example: $E(D)$ is the expected value of D.
- You can calculate the EV from the histogram:
  1. multiply each outcome by its probability,
  2. and sum up the terms.
- If all outcomes are equally likely, take the average.

- Think of the expected value as **the mean of the RV if infinitely repeatedly drawn**.
- *Average*, *Mean*, and *Expected* are often used interchangeably, even if not fully correct.
  - *Average* or *Mean* is for a specific (historical?) sample.
  - *Expected* is not known yet.
  - *Average* will be a realization of *Expected*.
  - On average, expect the *Average* to be the *Expected*.

10: Calculate E(RV)

What is the expected payoff of D?

11: Question

Is $E(D)$ a number or a RV?

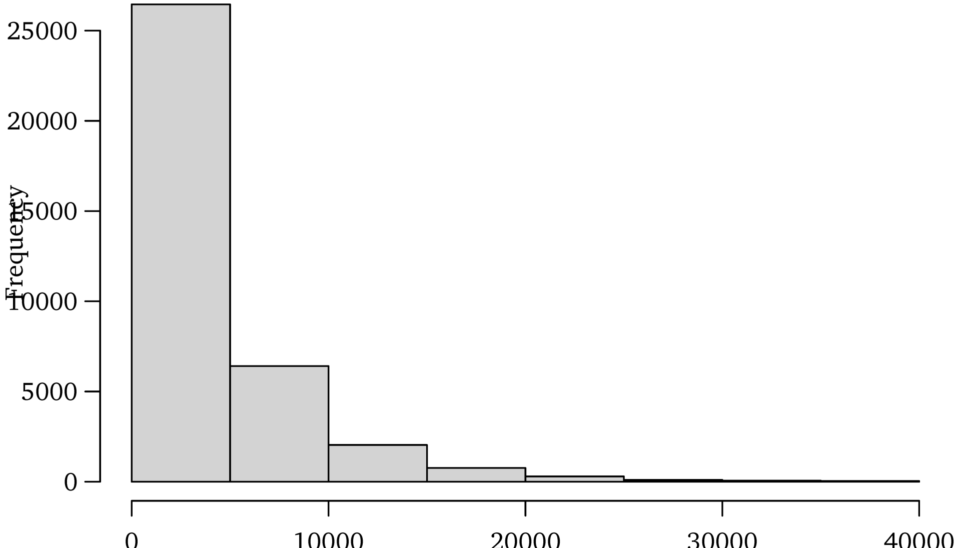
12: Is the Mean Mean?

- Is the expected payoff always the most likely outcome?
- Are half of all outcomes always below the mean?
- Can you have the mean outcome?
  - What is the average number of children per woman?

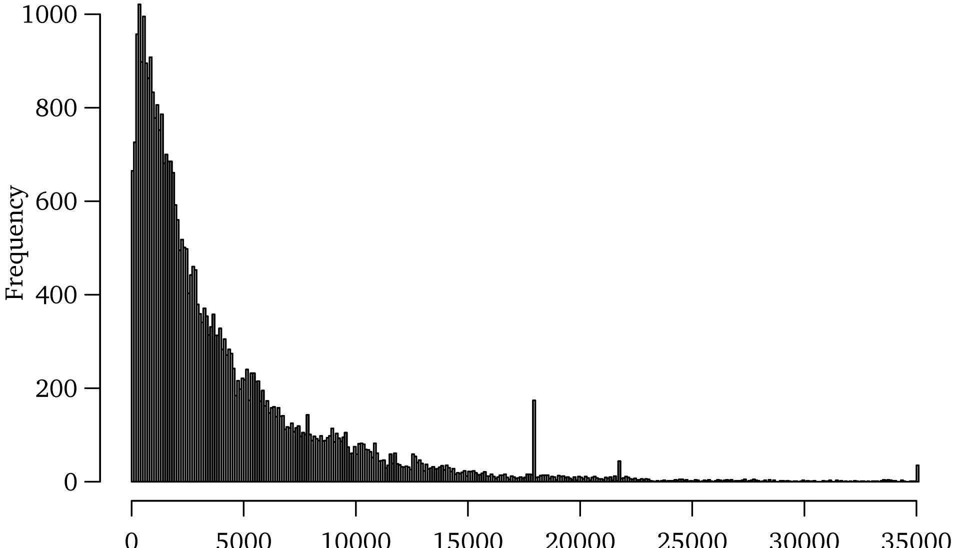
13: Central Statistics

- How Long Were Stocks Traded on US Public Exchanges?

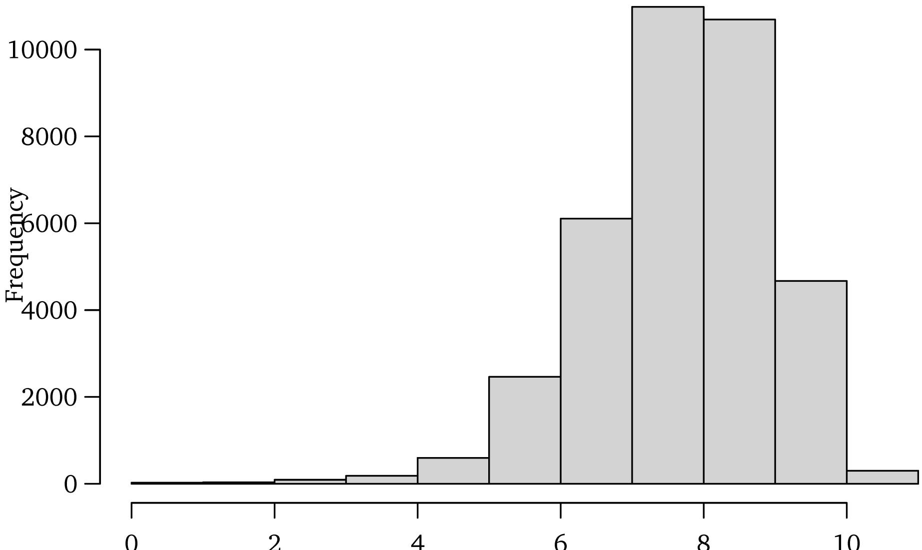
14: View 1: Mean = 3,992 days



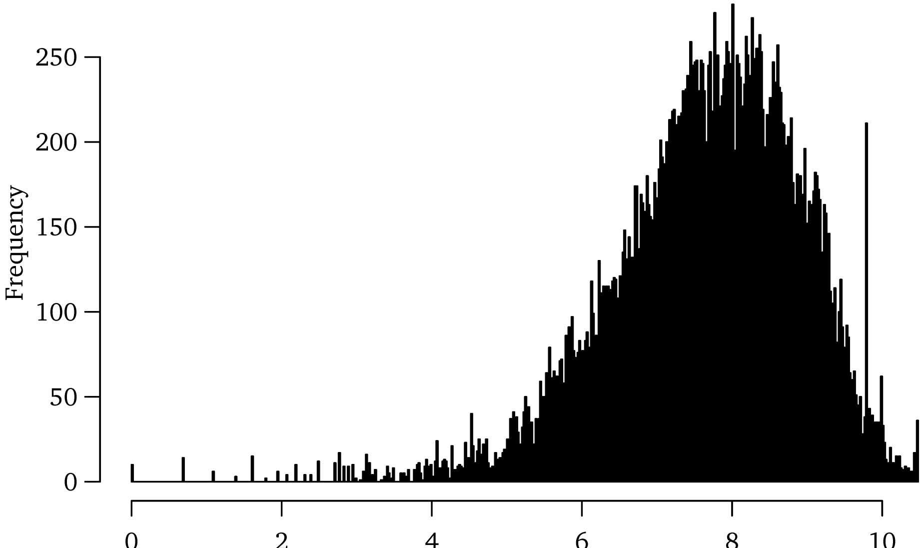
15: View 2: Mean = 3,992 days



16: View 3: Mean of Log Days = 7.679 days



17: View 4: Mean of Log Days = 7.679 days



18: WTH?

- The mean of days is 3,992d.
- The mean of $\log(days + 1)$ is 7.679.
- But exp(7.679) is equal to 2,162 days.

- **NOTE**: mean(log(days)) $\neq$ log(mean(days))?

19: Warning

- Histograms can be shaped quite differently, depending on the breaks

20: Different Central (Plus) Statistics

- What is the mean?
- What is the median?
- What is the mode?

- (What is the standard deviation?)

21: Different Central Statistics

- Mean: 3,992 days (10.9 years)
- Median: 2,425 days (6.6 years)
- Mode: 17,914 days (49 years)
  - why this high a number?
- Other Statistics:
  - SD= 4,515 (12.4y)
  - IQR= 1,037 – 5,308 (50-50 in vs out)
  - Min= 0. Max=35,064 (96 years).

22: Functions of RVs (Like Log(X))

Back to our Die:

- The $E(D)$ was \$78.
- What is the expected value of the die-squared, $E(D^2)$?
- Recall D
  - 1 = –\$6,
  - 2 = +\$36,
  - 3 = –\$12,
  - 4 to 6 = +\$150.

23: Linear and Non-Linear Functions of RVs

Is $E[D^2]$ the same as $[E(D)]^2$?

PS: Is $2 \times E(D) + 1 = E(2 \times D + 1)$

24: Measuring Unexpected?

What is an unexpected quantity?

25: Fair Bets

- What is *a* fair bet?
- What would it take for the *above* die-throw to become *a* fair bet?

26: Variance

- The variance is $Var(D) = E\{[D - E(D)]^2\}$
- Roughly speaking, the variance is the **expected squared deviation from the mean**.
  - This is just pseudo-intuitive.
  - Infinitely many draws: *Mean* of $(D - Mean(D))^2$.
    - the sample *turns into* the population variance.
  - The units on and meanings of variances are usually incomprehensible.
    - except that higher is more volatile
  - PS: Variance of Number of Days on CRSP = 20,385,225

27: Variance Example

Remember: E(D)=$78. From the RV's histogram, to obtain the variance: multiply each squared deviation from the mean by its probability, and add up terms.

State	Prob	Outcome	Oc–Mn	Squared
"1"	1/6	–$6	–$84	$$7,056
"2"	1/6	+$36	–$42	$$1,764
"3"	1/6	–$12	–$90	$$8,100
"4"-"6"	3/6	+$150	+$73	$$5,184
WMean:		+$78	$0	$$5,412

28: Standard Deviation (SD)

- The square-root of the variance:

$$\text{SD} = \sqrt{\$\$5{,}412} \approx \$73.57 \ .$$

- Think of the SD *as* the **typical deviation from the mean** of the next draw.
  - Not entirely correct, but close enough.
  - SDs have meaningful units.

29: Variance or SD?

- If the variance is higher, then the standard deviation is higher and vice-versa.
  - So, "more risk" is both "more VAR" and "more SD."
  - SD units are more meaningful,
  - but VAR is also commonly used.
  - Vanilla or Chocolate?

30: Very Nerdish: Signal Balance

- Let's say you want to combine two investment signals.
  - Think S_1 is (-100,+100) *and* S_2 is (-1,+1).
  - Your rank order would be
    - (S_1= 100, S_2= 1): 101
    - (S_1= 100, S_2= −1): 99
    - (S_1= −100, S_2= 1): -99
    - (S_1= −100, S_2= −1): −101
- Problem: All your best signals *are* really just S_1. S_2 is defacto unimportant.

- Solution: normalize signals first. Otherwise, your best signals may give too much weight to the signal that has the higher mean—or, if means *are* the same, the signal that has the higher variance.
- Two common normalizations:
  1. the "percentile rank"
  2. the "normalized value" (net of mean, divided by standard deviation), also called *a* "Z-value".

31: Big Leap of Faith

- What is the True Histogram?
  - Yes, we know the histogram for *a* die throw from physics.
  - But what is the histogram for the RoR on the stock market?
  - Therefore, we pretend that the (many) historical outcome realizations of RoRs can proxy for the true unknown histogram of RoRs.

32: Historical = Future?

- Then we pretend that this historical histogram applies to future RoRs, too.
- This is a "Hail Mary."
  - But once we assume we know the histogram, we can calculate expectations and do much more!
  - Employment program for finance professors?

33: Future Stock Returns as RVs

- This translation of the historical outcome histogram (distribution) into the future outcome histogram (distribution) is a *heroic* assumption.
- But what else can we do?
- Anyone have a better idea?
- Always remain mindful of the problem!

34: Expected Stock Returns

- Historical extrapolation generally works poorly for mean returns.
  - BTC was a very profitable investment in the past.
  - Will it also be in the future?
- Investors try to learn and adapt all the time.
- Moreover, there is *survivorship bias*.

35: Expected Market Returns

- What is the expected RoR on the overall stock market (say, the S&P 500 or DJ 30)?
  - Is it still the same that it was historically?
  - Why?

36: Risk of Stock Returns

- Extrapolation works well for variances, standard deviations, and market-betas (TBD).
  - Not perfect, but well.
  - Will show evidence soon.
  - Ergo, statistics works well for risk management!

37: Historical Example

- What are the $E(R)$ and $SD(R)$ if you only knew four historical stock returns:
  - +10%,
  - −5%,
  - +20%,
  - +15%.

38: Sample Historical Calculations

- Mean: 10%
- Each Deviation: 0%, 15%, 10%, 5%
- Squared Deviations: 0, 225, 100, and 25,
- Sum Squared Deviations: 350.
  - Now take the mean?!

39: Population vs. Sample: N vs N-1

- If these were true population probs, then divide this sum by N=4, which gives you 87.5.
  - This translates into an sd of 9.35%.
- If these were just historical samples, then divide sum by N-1=3, which gives you 116.67.
  - This translates into an sd of 10.8%.

40: Excel

- Excel *stdev()* uses N-1.
  - Excel assumes you knew only the sample.
- This makes little difference for large datasets.
  - But, it matters with *textN* = 4.
- (N vs $N-1$ matters only for small samples.)
  - Excel: *stdevp()* is not *stdev()*;
  - Excel: *varp()* is not *var()*.

41: Repeat Conceptual Warning

- Do not trust the historical means blindly for predicting future expected RoRs.
- For individual stocks (rather than big diversified portfolios), this would be exceedingly stupid.
- Even for big diversified portfolios (like the S&P 500), this is a big leap of faith.

42: BUT NOT ALL IS LOST

- (Recent) historical variances (and covariances and standard deviations) are usually good predictors of (short-term) future variances (and covariances and standard deviations).
- Advice: use 1–3 years of historical *daily* return data to estimate them.

43: Estimating Tail Risk?

- *Unfortunately*, a short time-series of historical numbers is usually not reliable enough to calculate tail-risk—the probability of a complete blow-up.
  - How do you estimate the risk of the next Space-X rocket exploding?
  - How do you estimate the prob of the next meteor hitting NYC from 80 years of data?
  - We would need more data than we have!
  - (Nerd Technical Alternative: Use Options.)
- Shrink them.

44: Monthly Stock Returns By Decade

Decade	Avg Ret	SD(Ret)
1990s	13%	13.4%
2000s	−4%	16.1%
2010s	13%	12.5%

45: Big Context Switch

Investor Preferences

46: Preferences over Bets

- If I offer you a bet of $+\$1$ if heads and $-\$1$ if tails, you pick a coin and someone else in class to throw it (at least 5 yards!), would you be willing to take this bet?
- If not, how much would I have to pay you?

47: Low Risk-Aversion

When is risk neutrality (or low risk-aversion) a good assumption?

48: Risk

Why do people climb mountains, play the lottery, drive motorcycles, or fly (Vans RV) airplanes?

49: Risk Preferences

- **Most of finance assumes that investors are risk-averse.**
- **When choosing between overall portfolios with equal expected returns, investors prefer the lower-risk one.**
- **Portfolio Risk can be measured by portfolio SD.**

50: Enough Theory

- now come applications of these concepts. . . !