Daniel Kahneman - Noise: Part II. Your Mind Is a Measuring Instrument

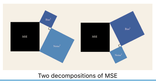
Part II. Your Mind Is a Measuring Instrument

- Judgment
  - A measurement in which the instrument is a human mind
    - Has implied goal of accuracy
  - not a synonym for thinking!
  - There is always some error
    - Bias
    - Noise
- Standard deviation
  - In statistics, the most common measure of variability
    - We will use it to measure noise in judgments

4. Matters of Judgment

- There can be some uncertainty about the answer (competent people might disagree)
  - But there is a limit to how much disagreement is admissible (expectation of bounded disagreement)
- The Experience of Judgment: An Example
  - An example of assessing Gambardi - a candidate for CEO position
    - 1. Selective attention to details
    - 2. Informal integration of memorized facts into overall impression
    - 3. Selecting a number from 0 to 100 to match impression - another source of variability
- What Judgment Aims to Achieve: The Internal Signal
  - Predictive judgment is Nonverifiable when
    - Case is fictitious
      - Gambardi does not exist
    - Answer is probabilistic (%)
      - 80%
    - The nature of professional judgment is nonverifiable
      - Underwriters will never know whether a particular policy was overpriced or underpriced
    - Forecast is conditional
      - If we go to war, we will be crushed
    - Forecast may be too long term
      - estimate of mean temperatures by the end of the 21st century
  - 👍 Internal signal of judgment completion
    - made you feel you got the judgment right
- How Judgment Is Evaluated: The Outcome and the Process
  - Verifiable judgments can be scored by
    - 👍 an objective observer on a simple measure of error
  - Both Verifiable and Nonverifiable judgments can be scored by
    - The same idea was in Michael Mauboussin's "More Than You Know" book
    - 👍 EVALUATING THE PROCESS OF JUDGMENT!
      - To observe how the process performs when it is applied to a large number of cases
        - The judgments are verifiable as an ensemble, although no single probability judgment can be declared right or wrong
      - Whether the process conforms to the principles of logic or probability theory
- Evaluative Judgments
  - involve multiple options and trade-offs between them
    - Example: Sentencing problem from Chapter 1
  - Evaluative judgment (selecting one decision) can be done after Predictive judgment (making few predictions)
  - boundary between predictive and evaluative judgments is fuzzy
- What's Wrong with Noise
  - Noise in Predictive judgements
    - Leads to serious consequences for those who rely on prediction
  - Noice in Evaluative judgements
    - Violates expectations of fairness and consistency
      - 👎 inconsistency damages the credibility of the system (organization)
- Undesirable but Measurable
  - noise is undesirable
  - 👍 But often measurable

5. Measuring Error

- If a manager most often predicts that projects will take half the time they ultimately take, and occasionally predicts they will take twice their actual time, it is unhelpful to say that the manager is "on average" right.
  - ⚠️ The different errors add up; they do not cancel out.
- Should GoodSell Reduce Noise?
  - Bias is simply the average of errors
- Mean Squares
  - mean squared error (MSE)
    - arithmetic mean as the value for which error is minimized
      - and it is supported by intuition
- The Error Equations
  - Error in a single measurement = Bias + Noisy Error
    - Bias = average error
    - + residual "noisy error"
      - The average of noisy errors is zero
    - Noise = standard deviation of measurements
      - = standard deviation of noisy errors
  - Overall Error (MSE) = Bias^2 + Noise^2
  - 
    - Two decompositions of MSE
  - 
    - Distribution of errors with bias reduced by half vs. noise reduced by half
  - 👍 => the benefit of reducing the Noise is the same as reducing the Bias
    - In this simplified model, we have assumed that noise and bias are equal
    - ❓ Q: Why benefits are the same? Bias could be much larger than Noise
  - ❓ Q: Why not calculate Noise as deviation. And WHEN we know the right answers we can calculate Average error = Bias and decide what is more important to reduce. Otherwise while reducing the Noise we can later find that Bias is much larger
- The Cost of Noise
  - Sometimes consequences of errors are asymmetrical
    - Underestimation is costly, but overestimation is catastrophic
    - Examples
      - Maximum load of evevator
      - When to leave to catch the train
  - 1. Predictive judgments
  - 2. Evaluative judgment
    - Where safety becomes a dominant consideration

6. The Analysis of Noise

- A Noise Audit of Sentencing
  - 16 cases (2 types, 6 parameters) were given to 208 judges
- Mean Sentences
  - There is no way to determine "just" sentence - so mean (average) sentence is considered as zero bias
    - However in reality this assumption is wrong due to racial discrimination and other reasons
- The Sentencing Lottery
  - When mean = 7.0 years, standard deviation = 3.8 years (quite large)
  - However in real life cases have much more parameters and the Noise is even higher!
- Some Judges Are Severe: Level Noise
  - Average level of each judge sentencing = personality trait
- Judges Differ: Pattern Noise
  - Pattern Errors = Pattern Noise
    - residual deviations
      - judge x case interaction
  - Patterned differences between judges in the influence of offense/offender characteristics
    - an additional form of sentence disparity
  - System Noise^2 = Level Noise^2 + Pattern Noise^2
- The Components of Noise
  - ⚠️ System Noise
    - undesirable variability in the judgments of the same case by multiple individuals
  - ⚠️ Level Noise
    - variability in the average level of judgments by different judges
  - ⚠️ Pattern Noise
    - variability in judges' responses to particular cases
    - contains some Occasion Noise
      - which can be treated as random error

7. Occasion Noise

- The Second Lottery
  - the professional's mood, the sequence of cases that are fresh in mind, and countless other features of the occasion
- Measuring Occasion Noise
  - Direct way to measure
    - Ask person to reevaluate the case
    - But hard to measure when cases are easily memorable
      - test-retest reliability
        - People tend to agree with themselves
  - Less direct way to measure
    - big data and econometric methods
- One Is a Crowd
  - 👍 wisdom-of-crowds effect
    - averaging the independent judgments of different people generally improves accuracy
  - 👍 the crowd within
    - can you get closer to the truth by combining two guesses from the same person
    - Second answer from the same person = 1/10 of second opinion from another person
  - 👍 Sleep on it, and think again in the morning
    - After three weeks of waiting, the second answer = 1/3 of second opinion from another person
    - ❓ Q: If second answer is better why should we still take the average? Why not just take the second answer?
  - 👍 Dialectical bootstrapping
    - instruction for second estimate
      - 1. Assume that your first estimate is off the mark.
      - 2. Think about a few reasons why that could be. Which assumptions and considerations could have been wrong?
      - 3. What do these new considerations imply? Was the first estimate rather too high or too low?
      - 4. Based on this new perspective, make a second, alternative estimate.
    - Accuracy of such two consecutive estimates = 1/2 of second opinion from another person
- Sources of Occasion Noise
  - What influence judgement
    - Mood
    - Stress and Fatigue
    - Weather influences Mood
    - ⚠️ After a series of decisions that go in the same direction, they are more likely to decide in the opposite direction than would be strictly justified
    - Order of cases
- Sizing Occasion Noise
  - 👍 Occasion Noise is smaller than differences among individuals
  - You are not always the same person
    - You are less consistent over time than you think
    - But you are more similar to yourself yesterday than you are to another person today
- Occasion Noise, Inner Causes
  - external influences (weather, distracting interventions, etc) gives only 11% of variability
  - ⚠️ Internal causes gives 89% of variability
    - e.g. basketball player muscles never execute exactly the same gesture
    - => occasion noise can NOT be eliminated!

8. How Groups Amplify Noise

- Group factors that may influence decisions (make noise)
  - who speaks first
  - who speaks last
  - who speaks with confidence
  - who is wearing black
  - who is seated next to whom
  - who smiles or frowns or gestures at the right moment
  - ...
- Noise in the Music
  - 👎 social influence
    - People like music tracks that were liked by others
    - creates significant noise
    - ⚠️ Popularity is self-reinforcing
    - TODO: Ensure to receive Positive first public feedback on my posts/talks
    - ⚠️ Who speaks first influences others!
    - Conformity effect!
- Beyond Music Downloads
  - political positions can be just like songs
  - Website
    - TODO: Immediately like first comments under my posts!
    - First vote up on the comment
      - The next viewer became 32% more likely to give an up vote
      - ❓ Idea to make Social Networks less noisy: hide first Votes and probably first Comments - at least for other Viewers
  - 👍 independence is a prerequisite for the wisdom of crowds
- Cascades
  - TODO: Collect expert views without cascade effects! Like in Scrum Poker
  - 👎 informational cascade
    - Happens when people speak one-by-one
    - Next person have more information about initially offered decision
    - People (if they don't have strong arguments) tend to agree with previous speakers
      - And it influences next people even more
  - 👎 social pressure cascades
    - People do not want to look disagreeable or silly
  - ⚠️ Unanimous decision can be bad sign that there was an effect of cascade!
    - ❓ Q: What about Sociocracy? They offer to speak in "circles" one by one and to make all decisions unanimously
- Group Polarization
  - 👎 when people in a group speak with one another, the group often end more unified, more confident, and more extreme
  - deliberating juries (who discuss their views of the case) were far noisier than statistical juries
    - TODO: Limit group discussions! However not clear what is the right limit? At least consider this effect!
    - ⚠️ Deliberation had the effect of increasing noise
    - ❓ WOW!!! Does it mean we should not discuss our views?