Initial Question

- Statistics graduate students from Berkeley made a substantial amount of money betting on roulette in Las Vegas.
- Were they simply lucky? If not, how did they do this and what did the casinos do in response?
The outcome of play is the bin a ball falls in as it circles a spinning wheel

- There are 38 possible bins
  - 00 and 0 colored green
  - 1, 2, ..., 36 odd are colored red, even are colored black
- Bet $1 - possible bets include
  - ball falls in a red bin - payoff is $2
  - ball falls in the set of bins 1-12 - payoff is $3
  - ball falls in one specific bin - payoff is $36

- Are the bets fair?

In Europe, the set of bins does not include 00, but the payouts are unchanged. Where would you rather play?
Assume that each of the outcomes \{00, 0, 1, \ldots, 36\} is equally likely

- Bet $1 on red (odd numbers) $2 payoff
- What is the expected payoff?
  - \(0 \cdot \frac{20}{38} + 2 \cdot \frac{18}{38} = \frac{36}{38} = 0.95\)
- Bet $1 on set \{1, \ldots, 12\} - $3 payoff; What is the expected payoff?
  - \(3 \cdot \frac{12}{38} = \frac{36}{38} = 0.95\)
- Bet $1 on set 5 - $36 payoff; What is the expected payoff?
  - \(36 \cdot \frac{1}{38} = \frac{36}{38} = 0.95\)
- Are you indifferent between these bets?
Recall: Payoff Variances

Assume that each of the outcomes \( \{00, 0, 1, \ldots, 36\} \) is equally likely

- What is the variance of the payoff from the bet on odd numbers?
  
  \[
  \text{Var}(P) = \mathbb{E}(P - \mathbb{E}P)^2
  \]
  
  \[
  (0 - .95)^2 \cdot \frac{20}{38} + (2 - .95)^2 \cdot \frac{18}{38} = 1.00
  \]

- Is the variance of the payoff from the bet on \( \{1, \ldots, 12\} \) higher than 1.00?

- What is the variance of the payoff from the bet on \( \{1, \ldots, 12\} \)?
  
  \[
  (0 - .95)^2 \cdot \frac{26}{38} + (3 - .95)^2 \cdot \frac{12}{38} = 1.94
  \]

- What is the variance of the payoff from the bet on 5?
  
  \[
  (0 - .95)^2 \cdot \frac{37}{38} + (36 - .95)^2 \cdot \frac{1}{38} = 33.21
  \]
If the center weight is perfectly balanced, all outcomes are equally likely.
If the center weight is not perfectly balanced, all outcomes are not equally likely.

- Suppose \( \Pr(\text{outcome} = 5) = \frac{1.5}{38} = 0.0395 \)

- Which bet would you place?
- What is the expected payoff from the bet on 5?
  - \( 36 \cdot \frac{1.5}{38} = \frac{54}{38} = 1.42 \)
  - a slight change in the probability leads to a large change in the expected payoff

- Would you place this bet?
Discretely observe \( n \) outcomes of roulette, indexed by \( i \), to estimate

\[
\P(\text{outcome} = 5) \equiv \rho
\]

\[
X_i = \begin{cases} 
1 & \text{if outcome } i \text{ is 5} \\
0 & \text{otherwise}
\end{cases}
\]

What is the mean of \( X_i \)?

\[
\E(X_i) = \P(\text{outcome} = 5)
\]

What is the variance of \( X_i \)?

\[
\Var(X_i) = (0 - \rho)^2 \cdot (1 - \rho) + (1 - \rho)^2 \cdot \rho
\]
\[
= \rho \cdot (1 - \rho) [\rho + (1 - \rho)] = \rho \cdot (1 - \rho)
\]

How would you estimate \( \rho \)?

by estimating \( \E(X_i) \) with \( \bar{X}_n \)
Estimator Properties

What happens to the behavior of our estimator as $n$ grows?

$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

- What is the mean of $\bar{X}_n$?
  - $E(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n \cdot \rho = \rho$
  - estimator unbiased even for $n = 1$

- What is the variance of $\bar{X}_n$?
  - $Var(\bar{X}_n) = \frac{1}{n^2} Var(\sum_{i=1}^{n} X_i)$
  - $= \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i)$ observations are independent
  - $= \frac{1}{n} \cdot \rho (1 - \rho)$

- As $n$ grows, variance declines to zero
  - with an arbitrarily large sample size, the variance is negligible
Definition: An estimator, $\bar{X}_n$, is consistent for a parameter, $\rho$, if for an arbitrarily large sample the estimator equals $\rho$ with probability 1

- For $\bar{X}_n$
  - $\text{Var} (\bar{X}_n) = \frac{1}{n}\rho (1 - \rho) \to 0$ as $n \to \infty$
  - with an arbitrarily large sample, no uncertainty
  - with an arbitrarily large sample, you know $\rho$

- $\bar{X}_n$ is a consistent estimator of $\rho$
The statistics graduate students discretely observed a large number of trials and constructed $\bar{X}_n$

- For the outcomes for which $\bar{X}_n$ exceeded $\frac{1}{38}$, bets were placed
  - the students made money, but not by random chance
- casinos now switch center weights at regular intervals
How Much is Enough?

How many outcomes should you observe before betting?

- As $n$ grows
  - variance of estimator declines (increasing accuracy)
  - you are more likely to be spotted and barred from the casino
Measurements of Accuracy

If center weight is balanced, \( \rho = \frac{1}{38} = 0.026 \)

Estimate \( \rho \) with \( \hat{\rho}_n = \bar{X}_n \)

\[ \mathbb{E}(\bar{X}_n) = \rho \quad \text{Var}(\bar{X}_n) = \rho \left(1 - \rho\right)/n \]

- 95 percent confidence interval based on

\[ (-1.96\sigma_{\bar{X}_n} \leq \bar{X}_n - \rho \leq 1.96\sigma_{\bar{X}_n}) \]

- How big should \( n \) be if we want 1% (.01) accuracy?
Necessary Sample Size

- How big should \( n \) be if we want our estimate to be within 1% of \( \rho \)?

\[
(-0.01 \leq \bar{X}_n - \rho \leq 0.01)
\]

- From the interval \((-1.96\sigma_{\bar{X}_n} \leq \bar{X}_n - \rho \leq 1.96\sigma_{\bar{X}_n})\)

  - \(1.96\hat{\sigma}_{\bar{X}_n} = 0.01\)
  - \(\sigma^2_{\bar{X}_n} = \rho (1 - \rho) / n\)

\[
n = \left(\frac{1.96}{0.01}\right)^2 \rho (1 - \rho)
\]

- If \( \rho \) is approximately .026

\[
n = 38,416 \times .025 = 973
\]
Increasing Accuracy

- How big should \( n \) be if we want our estimate to be within 0.1% of \( \rho \)?

\[
(-0.001 \leq \bar{X}_n - \rho \leq 0.001)
\]

- From the interval \((-1.96\sigma_{\bar{X}_n} \leq \bar{X}_n - \rho \leq 1.96\sigma_{\bar{X}_n})\)
  - \(1.96\hat{\sigma}_{\bar{X}_n} = 0.001\)
  - \(\sigma^2_{\bar{X}_n} = \rho (1 - \rho) / n\)

\[
n = \left( \frac{1.96}{0.001} \right)^2 \rho (1 - \rho)
\]

- If \( \rho \) is approximately 0.026

\[
n = 3,841,600 \times 0.025 = 97,285
\]

- accuracy increases by a factor of 10, required sample size increases by a factor of 100