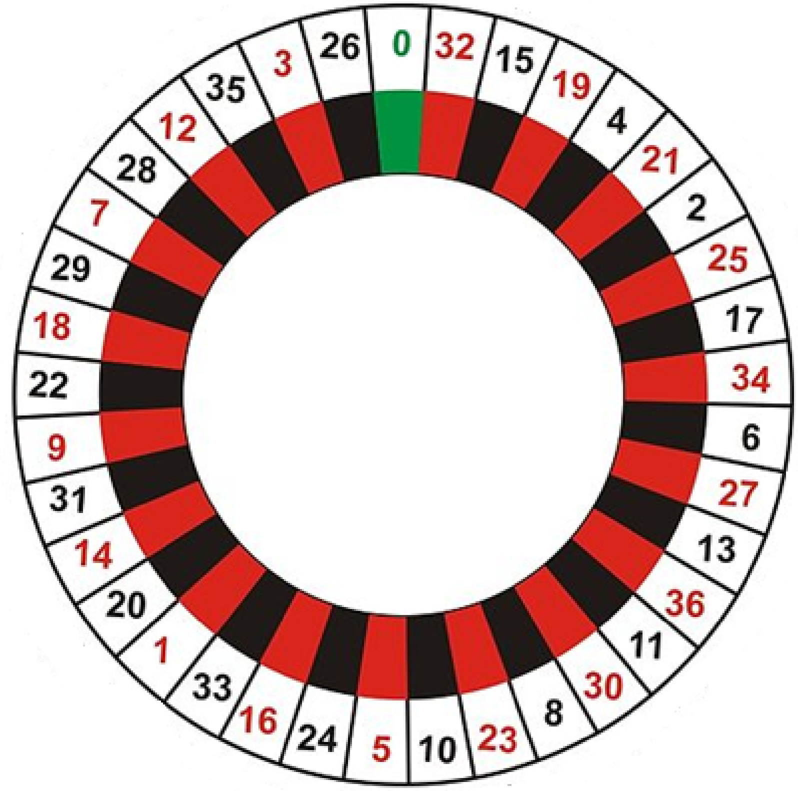
Inferential Statistics and Probability, Segment 2

Casinos Not in the Business of Being Fare

American Roulette



European Roulette

Comparing Different Games

Simulate betting a pocket for 20 trials of 100 spins each
Exp. return for Fair Roulette = 22.4%
Exp. return for European Roulette = -8.2%
Exp. return for American Roulette = 4.4%

Simulate betting a pocket for 20 trials of 1000 spins each
Exp. return for Fair Roulette = 3.68%
Exp. return for European Roulette = -5.5%
Exp. return for American Roulette = -4.24%

Simulate betting a pocket for 20 trials of 10000 spins each
Exp. return for Fair Roulette = 0.998%
Exp. return for European Roulette = -6.364%
Exp. return for American Roulette = -4.582%

Simulate betting a pocket for 20 trials of 100000 spins each
Exp. return for Fair Roulette = 0.125%
Exp. return for European Roulette = -3.313%
Exp. return for American Roulette = -5.5936%

Sampling Space of Possible Outcomes

- Never possible to guarantee perfect accuracy through sampling
- Not to say that an estimate is not precisely correct
- How many samples do we need to look at before we can have justified confidence on our answer?
  - Depends upon variability in underlying distribution

Quantifying Variation in Data

$$variance(X) = \frac{\sum_{x \epsilon X}(x-\mu)^2}{|X|}$$

$$\sigma(X) = \sqrt{\frac{1}{|X|}\sum_{x \in X}(x-\mu)^2}$$

- Standard deviation simply the square root of the variance
- Outliers can have a big effect
- Standard deviation should always be considered relative to mean

For Those Who Prefer Code

```
def getMeanAndStd(X):
    mean = sum(X)/float(len(X))
    tot = 0.0
    for x in X:
        tot += (x - mean)**2
    std = (tot/len(X))**0.5
    return mean, std
```

Confidence Levels and Intervals

- Instead of estimating an unknown parameter by a single value (e.g., the mean of a set of trials), a confidence interval provides a range that is likely to contain the unknown value and a confidence that the unknown value lays within that range
- “The return on betting on 2 twenty times in European roulette is -3.3%. The margin of error is +/- 1 percentage point with a 95% level of confidence.”

Empirical Rule

- Under some assumptions discussed later
  - ~68% of data within one standard deviation of mean
  - ~95% of data within 2 standard deviations of mean
  - ~99.7% of data within 3 standard deviations of mean

Applying Empirical Rule

```
numTrials = 20
resultDict = {}
games = (FairRoulette, EuRoulette, AmRoulette)
for G in games:
    resultDict[G().__str__()] = []
for numSpins in (100, 1000, 10000):
    print('\nSimulate betting a pocket for', numTrials,
          'trials of', numSpins, 'spins each')
    for G in games:
        pocketReturns = findPocketReturn(G(), 20, numSpins, False)
        mean, std = getMeanAndStd(pocketReturns)
        resultDict[G().__str__()].append((numSpins,
                                          100*mean, 100*std))
        print('Exp. return for', G(), '=', str(round(100*mean, 3))
              + '%,', '+/- ' + str(round(100*1.96*std, 3))
              + '% with 95% confidence')
```

Results

Simulate betting a pocket for 20 trials of 1000 spins each
Exp. return for Fair Roulette = 3.68%, +/- 27.189% with 95% confidence
Exp. return for European Roulette = -5.5%, +/- 35.042% with 95% confidence
Exp. return for American Roulette = -4.24%, +/- 26.494% with 95% confidence

Simulate betting a pocket for 20 trials of 100000 spins each
Exp. return for Fair Roulette = 0.125%, +/- 3.999% with 95% confidence
Exp. return for European Roulette = -3.313%, +/- 3.515% with 95% confidence
Exp. return for American Roulette = -5.594%, +/- 4.287% with 95% confidence

Simulate betting a pocket for 20 trials of 1000000 spins each
Exp. return for Fair Roulette = 0.012%, +/- 0.846% with 95% confidence
Exp. return for European Roulette = -2.679%, +/- 0.948% with 95% confidence
Exp. return for American Roulette = -5.176%, +/- 1.214% with 95% confidence