Due: March 1st, 2024, 5pm

Relative Frequency and Probability

Upload on Moodle code and Graph of problems marked to be submitted.

If $X \sim$ Bernoulli($\frac{1}{2}$). Suppose we toss a fair coin n times and let $X_i = 1$ if head occurs at the i-th toss and 0 otherwise. Let $A = \{1\}$ in the above, i.e. let

$$\hat{p}_n = \frac{|\{i : X_i = 1\}|}{n}.$$

1. Find $E[\hat{p}_n]$ and $Var(\hat{p}_n)$
2. Let us see via simulating in `R`, how does $\hat{p}_n$ approximate $\frac{1}{2}$
   (a) Find out by trying out examples of `x` what does the inbuilt function `cumsum`.
   ```
   > x = c(1,2,3,4)
   > cumsum(x)
   ```
   (b) Describe in two or three sentences what the `runningmean` function will output.
   ```
   > runningmean = function (x,N){
   + y = sample(x,N, replace=TRUE)
   + c = cumsum(y)
   + n = 1:N
   + c/n
   + }
   ```
   (c) Describe what you are observing from the plot:
   ```
   > u = runningmean(c(0,1), 1000)
   > v=1:1000; plot(u~v, type="l");
   ```
   (d) Using the `replicate` command on the same plot simulate the above experiment 10. First implement the code below
   ```
   > replicate(10, lines(runningmean(c(0,1), 1000)~v, type="l"))
   ```
   (e) *To be submitted:* Modify the above code to simulate 10 trials of the experiment (i.e. tossing a fair coin 1000 times).
   Extra credit: (5 points) color each plot with a randomly chosen `viridis` colour.
3. Using the above code(s) appropriately
   (a) Simulate 100 trials of the experiment (i.e. tossing a fair coin 1000 times) and calculate $\hat{p}_{1000}$ for each trial.
   (b) Plot the histogram of the `error`=$\hat{p}_{1000} - \frac{1}{2}$ (over the 100 trials).
   (c) *To be submitted:* Plot the histogram of the standarised error namely $\frac{\hat{p}_{1000} - \frac{1}{2}}{5}$ (over 100 trials).
   (d) What is your guess for the distribution for the standardised error ?