STAT 88: Lecture 24

Contents

Warm up: Draw 6 cards and count the number of red cards you get. To increase your odds of getting 3 red cards should you draw with or without replacement?

Last time

If $X \sim \text{HG}(N, G, n)$, then

$$\text{SD}(X) = \sqrt{n \cdot \frac{G}{N} \cdot \frac{N-G}{N}} \cdot \underbrace{\sqrt{\frac{N-n}{N-1}}}_{\text{fpc}}.$$

fpc < 1 if n > 1

If $Y \sim \text{Binomial}(n, \frac{G}{N})$, we have

$$\text{SD}(X) = \text{SD}(Y) \cdot \text{fpc}.$$

This implies that

$$\text{SD}(X) < \text{SD}(Y).$$

When $n \ll N$, $SD(X) \approx SD(Y)$

7.3. The Law of Averages

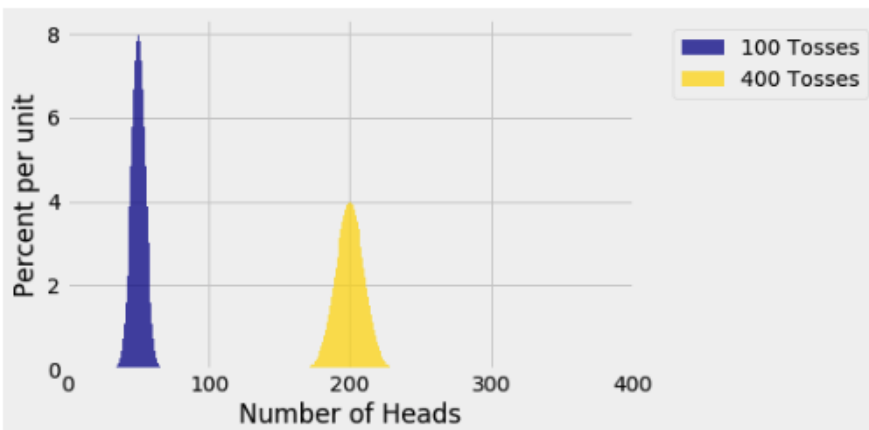
Informally, the law of averages is the familiar statement that if you toss a coin many times you get about half heads and half tails.

The Sample Sum

Let $X_1, \ldots, X_n$ be i.i.d. samples from a population with mean μ and SD σ, and let $S_n = X_1 + X_2 + \cdots + X_n$. As a running example, let the population distribution be Bernoulli distribution, i.e. you toss a fair coin n times and the samples are

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Bernoulli}\left(\frac{1}{2}\right).$$

Compare the distributions of S_{100} and S_{400}:



Observation: Both histograms have area 1. Yellow histogram is more spread out and hence $P(S_{400} = 200) < P(S_{100} = 50)$.

We can see this mathematically:

$$\begin{aligned}\text{Var}(S_n) &= \text{Var}(X_1 + \cdots + X_n) \\ &= \text{Var}(X_1) + \cdots + \text{Var}(X_n) \\ &= n\sigma^2,\end{aligned}$$

and so

$$\text{SD}(S_n) = \sqrt{n}\sigma.$$

The Sample Average

As before, let $X_1, \ldots, X_n$ be i.i.d. samples from a population with mean μ and SD σ, and let $S_n = X_1 + X_2 + \cdots + X_n$. Let $\bar{X}_n = S_n/n$ be the sample average.

We can get

$$E(\bar{X}_n) = \frac{1}{n}E(S_n) = \frac{1}{n}n\mu = \mu,$$
$$\text{SD}(\bar{X}_n) = \frac{1}{n}\text{SD}(S_n) = \frac{1}{n}\sqrt{n}\sigma = \frac{\sigma}{\sqrt{n}}.$$

The Square Root Law

In the language of estimation, the accuracy of an unbiased estimator can be measured by its SD: the smaller the SD, the more accurate the estimator.

For example,

$$\text{SD}(\bar{X}_{400}) = \frac{\sigma}{\sqrt{400}} = \frac{\sigma}{\sqrt{4}\sqrt{100}} = \frac{1}{2}\text{SD}(\bar{X}_{100}).$$

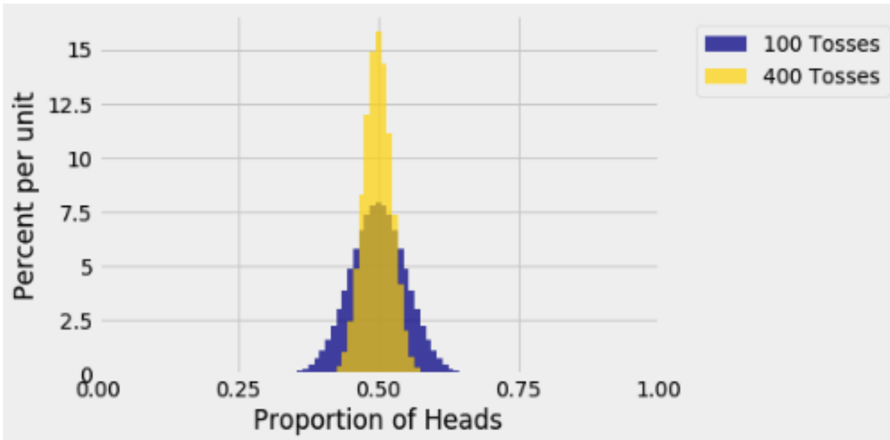
So $\bar{X}_{400}$ is *twice* as accurate as $\bar{X}_{100}$. For double the accuracy, we have to multiply the sample size by a factor of $2^2 = 4$.

If you multiply the sample size by a factor, the accuracy only goes up by the square root of the factor. This is called the **square root law**.

Concentration of Probabilities

You toss a fair coin n times. Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Bernoulli}\left(\frac{1}{2}\right)$.

Compare the distributions of $\bar{X}_{100}$ and $\bar{X}_{400}$:



Observation: Both histograms balance at 0.5. Yellow histogram is more concentrated around 0.5.

In general, the larger the sample size n, the more likely it is that the sample average $\bar{X}_n$ will be close to the population average μ.

Weak Law of Large Numbers: formally, for a fixed $c > 0$,

$$P(\mu - c < \bar{X}_n < \mu + c) = P(|\bar{X}_n - \mu| < c) \to 1 \text{ as } n \to \infty.$$

Interpretation: no matter how small c is, the chance that the sample mean is in the interval $(\mu - c, \mu + c)$ increases to 1 as the sample size grows.

The Law of Averages

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim}$ Bernoulli(p). Let $\bar{X}_n$ be the sample mean or sample proportion of 1s in your sample. The law of averages (=the law of large numbers) is: for each fixed $c > 0$,

$$P(p - c < \bar{X}_n < p + c) = P(|\bar{X}_n - p| < c) \to 1 \text{ as } n \to \infty.$$

Interpretation: when the sample size is large, the sample proportion of 1's is hugely likely to be in a small interval around p.

In a large number of rolls, it is hugely likely that the observed proportion of times the face appears is close to 1/6.