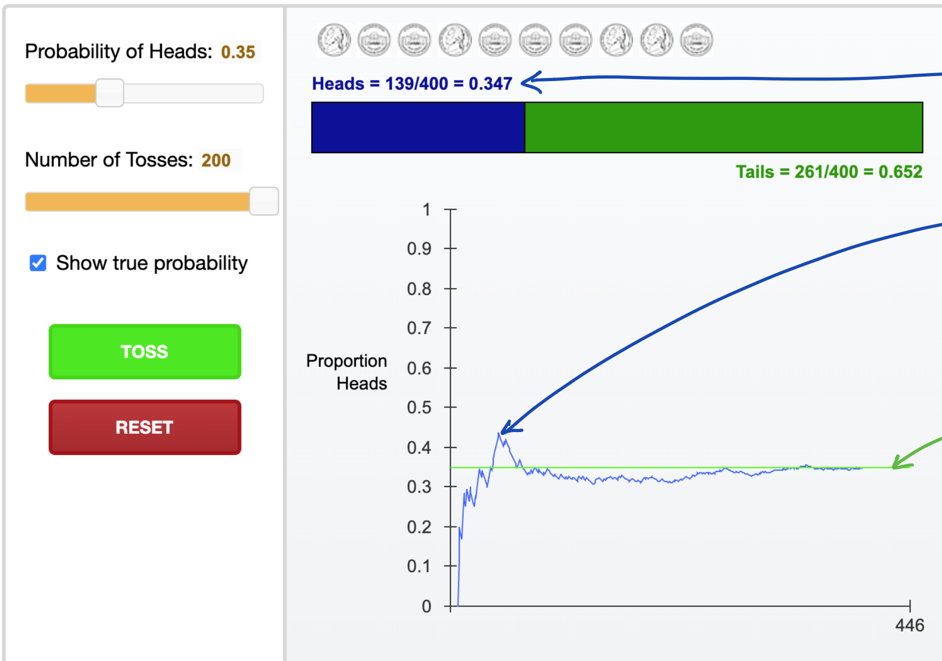
Weak law of large numbers

- The following statement seems correct by intuition: When the same experiment is repeated many times, the average value obtained over the experiments is close to the expectation of the experiment.
- Example: How do you know whether a given coin is fair?

- Another example: consider a Bernoulli random variable with probability of success (i.e., outcome is 1) p. If we conduct n Bernoulli trials then the average of the n outcomes will be very close to p (the expected value of the Bernoulli random variable) for large value of n.
- This phenomenon is called the weak law of large numbers.

Weak law of large numbers

- Example: If we toss a biased coin 400 times with probability of head = .35, it is very likely that approximately 140 times it will be head (i.e., success or outcome 1).



$\frac{\text{no. of Heads}}{\text{no. of trials}}$

average of the outcomes.

Expected value.

X is Bernoulli with $P(X=1) = .35$
$P(X=0) = .65$

$\Rightarrow E(X) = \sum x P(X=x)$
$= 0.35.$

Source: `http://digitalfirst.bfwpub.com/stats_applet/stats_applet_10_prob.html`

Visit this webpage to play with the applet.

Weak law of large numbers

- Let $X_1, X_2, \ldots,$ be a sequence of independent and identically distributed r.v.s, each having mean $E(X_i) = \mu$. Then, for any $\epsilon > 0$,

a random variable.

$$P\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) \to 0 \text{ as } n \to \infty.$$

Proof: Note that

$$E\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{1}{n}\left[E(X_1) + \ldots + E(X_n)\right] = \frac{1}{n} \cdot n\mu = \mu$$

($\because$ linearity of exp., recall Lecture 14)

$$Var\left(\frac{X_1 + \ldots + X_n}{n}\right) = \left(\frac{1}{n}\right)^2 \cdot Var(X_1 + \ldots + X_n)$$

($\because$ $Var(cX) = c^2 Var(X)$, recall Lecture 16)

Weak law of large numbers

$$= \frac{1}{n^2}\left[\mathrm{Var}(X_1) + \dots + \mathrm{Var}(X_n)\right] \quad (\because X_1, \dots, X_n \text{ are independent recall Lecture 16})$$

$$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Now, apply Chebyshev inequality:

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\mathrm{Var}\left(\frac{X_1 + \dots + X_n}{n}\right)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

Note that as $n \to \infty$, the R.H.S. $\frac{\sigma^2}{n\epsilon^2} \to 0$. Then, as $n \to \infty$,

$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| > \epsilon\right) \to 0$ since $P(\cdot)$ is lower bounded by 0 ($\because P(\cdot)$ is a non-neg. function) and upper bounded by $\frac{\sigma^2}{n\epsilon^2}$.

A result on normal distributions

- Before we study the central limit theorem, let's look at the relation between a normal distribution $\mathcal{N}(\mu, \sigma^2)$ and the standard normal distribution $\mathcal{N}(0, 1)$.
- Recall that if we want to solve problems on the standard normal distribution then we refer to its CDF table.
- What to do if we want to solve a problem on a normal distribution $\mathcal{N}(\mu, \sigma^2)$?
- Relation: If $Z \sim \mathcal{N}(0, 1)$, then $X = \mu + \sigma Z$ have the normal distribution $\mathcal{N}(\mu, \sigma^2)$. (proof: next slide) Partial proof.
- An application of this relation: note that $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$. Hence, the CDF of X can be found from the CDF of the standard normally distributed Z.

A result on normal distributions

- Relation: If $Z \sim \mathcal{N}(0,1)$, then $X = \mu + \sigma Z$ have the normal distribution $\mathcal{N}(\mu, \sigma^2)$.

<u>Partial</u> <u>Proof</u>: We will only prove that $E(X) = \mu$ and $Var(X) = \sigma^2$.

$$
\begin{aligned}
&E(X) \\
&= E(\mu + \sigma Z) \\
&= \underbrace{E(\mu) + \sigma E(Z)}_{\text{linearity of exp.}} \\
&= \mu + \sigma \cdot 0 \\
&= \mu
\end{aligned}
$$

recall lecture 14

$$
\begin{aligned}
&Var(X) \\
&= Var(\mu + \sigma Z) \\
&= Var(\sigma Z) \quad \because Var(X+c) = Var(X) \\
&= \sigma^2 \underbrace{Var(Z)}_{1} \quad \because Var(cX) = c^2 Var(X) \\
&= \sigma^2
\end{aligned}
$$

recall lecture 16

recall lecture 16

A result on normal distributions

- Example (application of the relation): Let $X \sim \underbrace{\mathcal{N}(\overset{\mu}{-1}, \overset{\sigma^2}{4})}$. What is $P(|X| < 3)$?

normal but <u>not</u> standard normal.

- We know that if $X \sim \mathcal{N}(\mu, \sigma^2)$ then $X = \mu + \sigma Z$ where $Z \sim \mathcal{N}(0,1)$. $\Rightarrow Z = \frac{X-\mu}{\sigma} \Rightarrow Z = \frac{X+1}{2}$.

Now,

$$
\begin{aligned}
P(|X| < 3) &= P(-3 < X < 3) \\
&= P\left(\frac{-3+1}{2} < \frac{X+1}{2} < \frac{3+1}{2}\right) \\
&= P(-1 < Z < 2) \\
&= F_Z(2) - F_Z(-1) \\
&= .9772 - 0.1587 = 0.8185
\end{aligned}
$$

($Z \sim \mathcal{N}(0,1)$ $\Rightarrow$ use the CDF table)