Example: Suppose Y is a Binomial random variable with $n = 5000$ and $p = \frac{1}{10}$. Find $P(Y \le 520)$. We can use CLT to find a good approximation to this probability. The exact value, by the way, is

$$\sum_{y=0}^{520} \binom{1000}{y} \left(\frac{1}{10}\right)^y \left(\frac{9}{10}\right)^{1000-y}$$

← 521 terms here! Difficult to calculate. CLT approximation is more feasible to actually carry out.

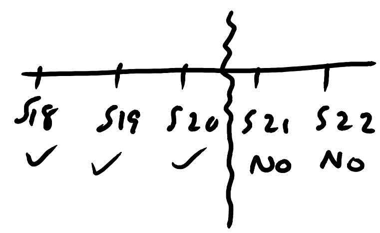
CLT is applicable because Y has the same distribution as $X_1 + X_2 + \dots + X_{1000}$ where the X_j's are independent Bernoulli random variables, each with $p = \frac{1}{10}$. $E(X_j) = \frac{1}{10}$, $Var(X_j) = (\frac{1}{10})(\frac{9}{10})$.

Continuity Correction

$P(Y \le 520)$

↑

Y is integer valued



$P(X \le 520.9)$

$P(X \le 520.1)$

$P(X \le 520.33)$

$P(X \le 520.79)$

We split the difference and use $P(Y \le 520.5)$

$$= P\left(\frac{Y - (5000)(1/10)}{\sqrt{5000(1/10)(9/10)}} \le \frac{520.5 - (5000)(1/10)}{\sqrt{5000(1/10)(9/10)}} \right)$$

$$\approx P(Z \le 0.97) = 0.8340.$$

We are using a continuous random variable to approximate the behavior of a discrete random variable.

We chose 520.5 because any cutoff from 520.0 to 520.999..... Would be OK and we want to avoid error in rounding as much as possible.