Estimating A Population Proportion

A roulette wheel at a local casino has 38 slots: 18 each of red and black and 2 neutral. In a test of the fairness of the wheel, red came up 8 times in 10 spins of the wheel. Is this convincing evidence that the wheel is unfair?

The number of spins, y, coming up red will have a $b(10, p)$ distribution, where p is the chance the wheel comes up red. To test the fairness of the wheel, we can test the statistical hypotheses $H_0 : p = 18/38$ versus $H_a : p \neq 18/38$. Under H_0, y has a $b(10, 18/38)$ distribution. The values for $f(y) = 10y(18/38)^y(1 - 18/38)^{(10-y)}$ are the following:

y	$f(y)$	y	$f(y)$
0	0.00163	6	0.18203
1	0.01468	7	0.09361
2	0.05945	8	0.03159
3	0.14268	9	0.00632
4	0.22473	10	0.00057
5	0.24270		

The observed value of y is $y^* = 8$, so the p-value is the sum of all values of $f(y)$ which are less than or equal to $f(y^*) = 0.03159$: $0.00163 + 0.01468 + 0.03159 + 0.00632 + 0.00057 = 0.05479$. This p-value is greater than 0.05 and leads us to not reject H_0 at the usual 0.05 level of significance.

Someone looking to find the wheel unfair might, after seeing the data change the alternative hypothesis to $H_a : p > 18/38$. The p-value would then be sum of the $f(y)$ values for $y \geq y^* = 8$: $0.03159 + 0.00632 + 0.00057 = 0.03848$, which leads to the rejection of H_0 at the 0.05 significance level and gives the conclusion that the wheel is unfair. However, letting the data decide the hypotheses is called data snooping and is unethical. The hypotheses should always be specified and the test procedure decided prior to viewing the data.