We want to test the hypothesis that more than 30% of U.S. households have Internet access (with a significance level of 5%). We collect a sample of 150 households and find that 57 have access.

(1) 假设：

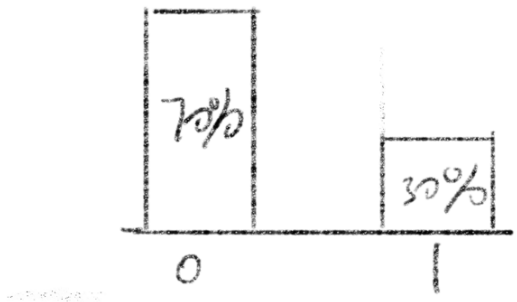
$$H_0 : P \leq 30\%$$
$$H_1 : P > 30\%$$

(2) 求均值

如果零假设成立，接入互联网最大的概率为30%：

$$P_{H_0} = 0.3$$

则分布如下：



这是个伯努利分布。可以求得方差与标准差：

$$\sigma_{H_0}^2 = P_{H0}(1 - P_{H0}) = 0.3 \times 0.7 = 0.21$$
$$\sigma_{H_0} = \sqrt{0.21} = 0.46$$

(3) 样本的统计量

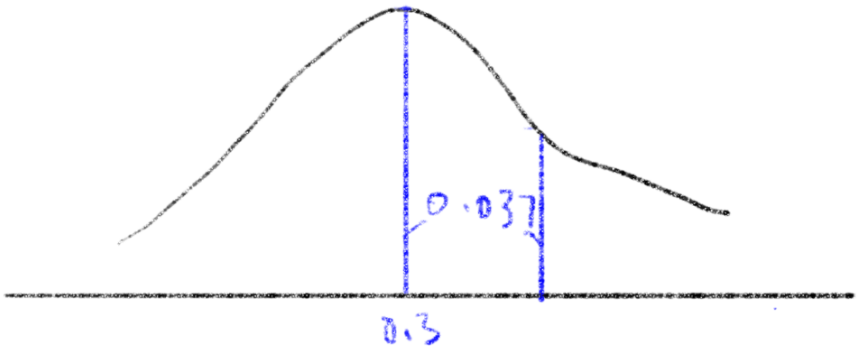
求样本的概率：

$$\bar{P} = \frac{57}{150} = 0.38$$

(4) 抽样分布的统计量

零假设成立，则抽样分布统计量如下：

$$\mu_{\bar{P}} = 0.3$$
$$\sigma_{\bar{P}} = \frac{\sigma}{\sqrt{150}} = \frac{0.46}{12.25} = 0.037$$



(5) 置信度

我们要判断抽样的概率在零假设误差为5%的置信程度，也就是看抽样的概率是否在抽样分布的95%置信区间内。

查z分布表，95%的置信度的z统计量是1.65。而我们的样本的z统计量为：

$$Z = \frac{0.38 - 0.3}{0.037} = 2.14$$

样本的z值2.14大于p值（极端值）1.65。所以倾向于拒绝零假设