

Задача 2

1. Дима собрал 3 наблюдения. Значения выборки: [1, 1, 0]. Посчитаем с помощью numpy среднее значение и дисперсию:

```
1 a = np.array([1, 1, 0])
2 print('Среднее:', a.mean())
3 print('Дисперсия:', a.std() ** 2)
```

[13] ✓ 0.0s

... Среднее: 0.6666666666666666
Дисперсия: 0.2222222222222222

В 2/3 случаев булочки были в наличии

2. $P(A) = p * (1 - p)$.

Так как мы считаем, что оба события должны произойти (булочки есть, булочек нет), то мы должны перемножить вероятности. Так как p - вероятность того, что булочка есть $\Rightarrow (1 - p)$ - вероятность того, что булочки нет

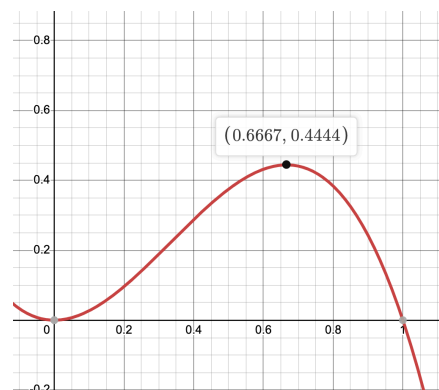
Рассмотрим два случая: Булочки есть-булочек нет И булочек нет-булочки есть. Следовательно, у нас два идентичных случая ($P(A) = p * (1 - p) = (1 - p) * p$) $\Rightarrow$ Общая вероятность равна $P(A) = 2 * p * (1 - p)$

3. Так как не сказано, в каких именно по счету днях булочки были/не было $\Rightarrow P(A) = p * p * (1 - p) + p * (1 - p) * p + (1 - p) * p * p = 3 * p * p * (1 - p) = 3 * p^2 * (1 - p)$

Итого, вероятность повторить наблюдения как у Димы равна $P(A) = 3 * p^2 * (1 - p)$

На оси OY отмечена вероятность, на OX - значение p . Видим, что максимум достигается, когда $p \approx 0.44$.

Так как вероятность $P(A)$ при $p = 0.44$ примерно равна 2/3, что равно доле случаев, когда булочки были в наличии, можно сказать, что это значение p - это действительно вероятность того, что в дни, рассматриваемые Димой в выборке, булочки могли быть испечены с вероятностью $p = 0.44$ (44%).



5. Нет, так как выборка, по которой мы считали вероятность, слишком мала, и нельзя предположить, что это значение останется таким же при рассмотрении репрезентативной выборки.

6. Гипотеза: Более популярная и расположенная в более людном месте пекарня производит булочки с повидлом в большем числе дней.

Чтобы проверить ее, Диме необходимо изучить проходимость мест, в которых расположены обе пекарни. Также необходимо узнать, не является ли пекарня с большим числом произведенных булочек с повидлом популярной франшизой и люди специально едут туда попробовать их фирменные булочки с повидлом, а также широту ассортимента пекарен.

Таким образом, ему надо собрать данные о:

- числе людей, проходящих мимо этой пекарни (проходимость улицы)
- доле людей, которые зашли в пекарню
- ассортимент (посмотреть, если пекарня с большим числом булочек с повидлом производит только булочки с повидлом и булочки с маком, например, а у другой пекарни более широкий ассортимент выпечки)