Задача: Подкинули монету N раз. Кол-во случаев, когда выпал орёл, на 10% больше, чем кол-во случаев, когда выпала решка. При каком N мы можем сказать, что монета «нечестная» (орёл и решка выпадают с разной вероятностью)?

Решение: Проверим гипотезу H_0, гласящую, что вероятность выпадения орла $P_{\text{орел}}$ равна вероятности выпадения решки $P_{\text{решка}}$ ($P_{\text{орел}} = P_{\text{решка}} = 0.5$). Для этого воспользуемся критерием согласия Пирсона. Сразу выберем уровень значимости α – пусть он равен 0.1. (Если P value меньше α, отклоняем нулевую гипотезу. P value – вероятность наблюдать полученные или еще более экстремальные результаты при условии, что нулевая гипотеза верна. P value также показывает шанс неверно отклонить H_0)

Итак, если монетка честная, мы ожидаем увидеть $\frac{N}{2}$ выпадения орла и столько же решек. По условию задачи мы видим $\frac{11}{21}N$ орлов и $\frac{10}{21}N$ решек.

	Орел	Решка
Наблюдаемое	$\frac{11}{21}N$	$\frac{10}{21}N$
Ожидаемое	$\frac{N}{2}$	$\frac{N}{2}$

Подсчитаем $\chi^2 = \sum \frac{(\text{наблюдаемое}-\text{ожидаемое})^2}{\text{ожидаемое}}$, суммируем по всем случаям (по орлам и решкам)

$$\chi^2 = \left(\frac{\left(\frac{22}{42}N - \frac{21}{42}N\right)^2}{\frac{1}{2}N}\right) + \left(\frac{\left(\frac{20}{42}N - \frac{21}{42}N\right)^2}{\frac{1}{2}N}\right) = \frac{1}{882}N + \frac{1}{882}N$$

$$\chi^2 = \frac{1}{441}N$$

Выяснили, как значение критерия зависит от количества бросков. Заглянем в таблицу и узнаем критическое значения для выбранного уровня значимости и нашего числа степени свобод (1 степень свободы в данном случае).

$$\chi^2_{\text{крит.}} = 2.71$$

Если значение критерия превышает критическое, можем заявить с уверенность 90%, что монетка нечестная. Установим, при каких значениях числа бросков это происходит.

$$\frac{1}{441}N \geq 2.71$$

$$N \geq 1195.11$$

Таким образом, если наблюдаем описанную ситуацию при 1196 и более бросках монетки, монету можно назвать нечестной. Проверим полученный результат с помощью вычислений в Python.

```
>>> from scipy.stats import chisquare as c
>>> c([626.47619, 569.52381])
Power_divergenceResult(statistic=2.712018049886617, pvalue=0.09959492826549766)
```

Здесь мы провели тест с помощью SciPy, введя в качестве наблюдаемых значений 11/21 и 10/21 от 1196. Тест показал критический уровень статистики и pvalue 0.1, на которое мы и ориентировались. Попробуем сделать чуть меньше бросков – например, просто 1190. Введем соответствующие значения:

```
>>> c([623.3333, 566.6667])
Power_divergenceResult(statistic=2.6984063492100856, pvalue=0.10044860652933014)
```

А вот здесь pvalue хоть немного, но больше уровня значимости – нулевую гипотезу отклонить не можем.

Если сделать больше, бросков, например, 1500, pvalue упадет еще сильнее, а мы еще больше убедимся в нечестности монетки:

```
>>> c([785.7, 714.3])
Power_divergenceResult(statistic=3.3986400000000083, pvalue=0.06525019650344582)
```

Вообще, строго говоря, при 1196 подбрасываниях мы не сможем получить ровно на 10% больше случаев орла – 1196 не делится на 21. А вот 1197 делится – так что ответом будет 1197 и более.