Handout: Test for independence

A survey was conducted to evaluate the effectiveness of a new flu vaccine that had been administered in a small community. It consists of a two–shot sequence in two weeks.
A survey of 1000 residents the following spring provided the following information:

	No vaccine	One Shot	Two Shots	Total
Flu	24	9	13	46
No Flu	289	100	565	954
Total	313	109	578	1000

1. Hyp.: H_0 : Getting the flu and the vaccine are independent versus H_a : H_0 is not true, $\alpha = 0.05$

2. Assumptions: random samples, $\{\hat{\mu}_{ij} \geq 5\}$ (check later).

3. Test Statistic: X_0^2 or G_0^2, $df = (I-1)(J-1)$

   First find the expected cell counts:

	No vaccine	One Shot	Two Shots	Total
Flu	24	9	13	46
	14.40	5.01	26.59	
No Flu	289	100	565	954
	298.60	103.99	551.41	
Total	313	109	578	1000

   Assumptions are met.
   $df = (3-1)(2-1) = 2$ and

$$X_0^2 = \sum_{i,j} \frac{(n_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}} = 6.404 + 3.169 + 6.944 + 0.309 + 0.153 + 0.335 = 17.31$$

   and

$$G_0^2 = 2 \sum_{ij} n_{ij} \ln \frac{n_{ij}}{\hat{\mu}_{ij}} = 2\left(24 \ln \frac{24}{14.40} + \cdots + 565 \ln \frac{565}{551.41}\right) = 17.26$$

4. P-value: P-value$=P(\chi^2 > X_0^2) < 0.005$ or P-value$=P(\chi^2 > G_0^2) < 0.005$ (table VII).

5. Reject H_0, since the P-value is smaller than $\alpha = 0.05$

6. At significance level of 5% the data provide sufficient evidence that there is an association between the vaccination status and getting the flu.

Standardized cell residual:

$$Z_{ij} = \frac{n_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}(1 - p_{i+})(1 - p_{+j})}}$$

is the standardized cell residuals for the cell in row i and column j.

If H_0 would be true these would be approximately standard normally distributed. There fore standardized cell residuals which do not fall between $-1,96$ and 1.96 indicate that the particular cell gives evidence that H_0 is violated.
Should we do a multiple comparison?

The standardized cell residuals for the flu example:

	No vaccine	One Shot	Two Shots	Total
Flu	24	9	13	46
	3.12	1.93	-4.15	
No Flu	289	100	565	954
	-3.12	-1.93	4.15	
Total	313	109	578	1000

The sample odds ratio for no vaccine and 2 shots is

$$\hat{\theta} = \frac{24 \times 565}{289 \times 13} = 3.6$$

The estimated odds for getting the flu are 3.6 times higher for people who did not get vaccinated compared to those who got 2 shots. Or we can say, that the estimated odds for getting the flu are 260% higher for people who did not get vaccinated compared to those who got 2 shots.
Yep, the vaccine works.