2014 Q1

(a) on campus: $\frac{17+7}{33} = .7273$ (b) off campus: $\frac{25+12}{67} = .5522$

(b) comparing between residential status for each number of activities:

	on campus	off campus	(difference)
no activities	$\frac{9}{33} = .27$	$30/67 = .45$	$.45 - .27 = .18$
one activity	$17/33 = .52$	$25/67 = .37$	$.37 - .52 = -.15$
2+ activities	$7/33 = .21$	$12/67 = .18$	$.18 - .21 = -.03$

For those participating in 2+ activities, the proportions are similar for both residential statuses (3% difference), but for no activities off campus is substantially higher (18% difference) and for 1 activity on campus is substantially higher (15% difference). With differences this large we conclude participation in extracurricular activities is <u>not independent</u> of residential status (there <u>is</u> an association).

(c) Hypotheses are given, with a p-value = .23:
with $\alpha = .05$, p-value = .23 is high so we fail to reject H_0.
We <u>do not</u> have sufficient statistical evidence to conclude that there is an association between participation in extracurricular activities and residential status.

*Note: are (b) and (c) contradictory?
<u>No</u>: part b is analyzing the <u>sample</u> there is a measurable difference in the sample, and according to our loose "rule of thumb" if any category difference is 15% or higher it is probably true that there is an association.
But the inference test which generates a p-value is the formal, more correct way to determine things about the <u>population</u>. There were fairly large differences in this sample, but (considering the sample size) these differences aren't large enough to assume an association for the entire population.