Estimating The Difference of Two Population Proportions

In a recent survey on academic dishonesty simple random samples of 200 female and 100 male college students were taken. 24 of females and 26 of the males agreed or strongly agreed with the statement "Under some circumstances academic dishonesty is justified." Researchers would like to compare the population proportions, p_f of all female and p_m of all male college students who agree or strongly agree with this statement. The statistical hypotheses they want to test are:

$$\begin{array}{lrcl} H_0: & p_f - p_m & = & 0 \\ H_a: & p_f - p_m & \neq & 0 \end{array}$$

The point estimate of $p_f - p_m$ is

$$\hat{p}_f - \hat{p}_m = 24/200 - 26/100 = -0.140,$$

and the estimate of the common value of p_f and p_m under H_0 is $\hat{p} = (26 + 24)/(200 + 100) = 0.167$.

Thus, the observed value of the test statistic is

$$z^* = \frac{24/200 - 26/100}{\sqrt{(0.1\overline{66})(0.8\overline{33})\left(\frac{1}{200} + \frac{1}{100}\right)}} = -3.067.$$

Since $y_f = 24$, $200 - y_f = 176$, $y_m = 26$, and $100 - y_m = 74$ all exceed 10, we may use the normal approximation:

$$p_- = P(N(0,1) \leq -3.04) = 0.0011, \; p^+ = P(N(0,1) \geq -3.04) = 0.9989,$$

and

$p\pm = 2\min(0.9989, 0.0011) = 0.0022$,

this last being the p-value we want. Such a small p-value leads to rejection of H_0 in favor of H_a, and the researchers conclude p_f and p_m are different.