$$z=\frac{\overline{P}_1-\overline{P}_2}{\sqrt{p_1 q_1(\frac{1}{n_1}+\frac{1}{n_2})}}$$

ここで,いま,の二つ母集団からそれぞれ実際に抽出された任意標本の比率を $\overline{p}_1, \overline{p}_2$ としますと, $\overline{p}_1=\frac{r_1}{n_1}$ $\overline{p}_2=\frac{r_2}{n_2}$ で, $z=\frac{\overline{p}_1-\overline{p}_2}{\sqrt{p_1 q_1(\frac{1}{n_1}+\frac{1}{n_2})}}$………(※)

は，標準正規分布から，任意に抽出された1つの数値と考えられます。

ところで，この場合，母比率 $p_1(=p_2)$は不明ですから，この zの値は計算不可能です。それで， n_1, n_2 がどちらも大のとき，目安としてどちらも 100以上のとき， $p_1(=p_2)$の代わりに $p=\frac{r_1+r_2}{n_1+n_2}$を代用して検定しようというのですから，結局，大きさ n_1+n_2の標本のうち，目当てのものが r_1+r_2個だけ出現したわけですから， $p=\frac{r_1+r_2}{n_1+n_2}$を $p_1(=p_2)$の推定値として 用いるのは当然のことでしょう。また， $r_1=n_1\overline{p}_1$ ， $r_2=n_2\overline{p}_2$ ですから，

$p=\frac{n_1\overline{p}_1+n_2\overline{p}_2}{n_1+n_2}$とも表せます。もちろん， $q=1-p$で，

上式(※)は $$z=\frac{\overline{p}_1-\overline{p}_2}{\sqrt{pq\left(\frac{n_1+n_2}{n_1 n_2}\right)}}$$

この両辺の絶対値をとったものが，〔7〕の2の$|z|$の式なのです。

(例17) 次の表は，昭和45年度，53年度に，それぞれ県下の6年生から，任意標本を抽出して実施した算数のテストの正答率の一部である。

年度 / 標本数 / 小問番号	45	53
	887	1110
1	68.1	71.5
2	80.0	84.3

年度によって，県下6年生全体の正答率に差があるといってよいか。

(解) ○小問1について，

1．仮説 H_0: $p_1=p_2$,年度によって正答率の差はない，とします。

　対立仮説は H_1: $p_1 \neq p_2$

2． $p=\frac{n_1\overline{p}_1+n_2\overline{p}_2}{n_1+n_2}=\frac{887\times 0.681+1110\times 0.715}{887+1100 \atop 1997}=0.700$