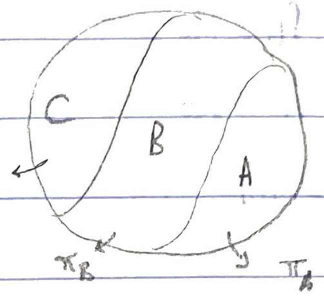
Chapter 12: Tests of Goodness of fit and Independence.

12.1: Goodness of fit and Independence.



★ In general:

→ $H_0: \pi_1 = \pi_{1_0}, \ \pi_2 = \pi_{20}, \ \dots, \ \pi_K = \pi_{K0}$.

H_1: The pop. propotion are not $\pi_1 = \pi_{1_0}, \dots \pi_K = \pi_{K0}$.

→ π_i: Population propotion of category i.

π_{i0}: Hypothesied value of the pop. of category i.

$i = 1, \dots, K$.

→ n: sample size.

f_i: observed frequencies.

e_i: expeted frequencies.

$$\boxed{e_i = n\,\pi_{i0}} \qquad \boxed{\sum_{i=1}^{K} f_i = \sum_{i=1}^{K} e_i = n}$$

→ Test statistic:

$$\chi^2 = \sum_{i=1}^{K} \frac{(f_i - e_i)^2}{e_i}$$

→ Reject H_0 if p-value $\le \alpha$

Reject H_0 if $\chi^2 \ge \chi^2_\alpha$.

$\left\{ \begin{array}{l} df = K-1 \\ e_i \ge 5 \ \forall i. \end{array} \right.$

H_0 : $\pi_A = 0.40$, $\pi_B = 0.40$, $\pi_C = 0.20$.

H_1 : The pop proportions are not $\pi_A = 0.40$, $\pi_B = 0.40$, $\pi_C = 0.20$.

Q1 :

$K = 3$ (A B C)

$n = 200$	H_0	H_1
$f_A = 60$	$e_A = 80$	$\pi_A = 0.40$
$f_B = 120$	$e_B = 80$	$\pi_B = 0.40$
$f_C = 20$	$e_C = 40$	$\pi_C = 0.20$

f مجموعهم لازم يساوي n

$\alpha = 0.01$.

$$\rightsquigarrow \chi^2 = \frac{(60-80)^2}{80} + \frac{(120-80)^2}{80} + \frac{(20-40)^2}{40}$$

يستخدم جدول كاي سكوير

$\boxed{\chi^2 = 35}$ $\quad$ $\star$ $df = K-1 = 3-1 = 2$

$e_i \geq 5 \quad \forall i$.

a. P-value

من جدول كاي سكوير بدور وين بلاقي χ^2 مع df

df	0.01	0.005
2	9.210	10.597

↓ upper tail area .

↑ $\chi^2 = 35$

اعلى قيمة فالجدول

$\Rightarrow$ upper tail area < 0.005

$\rightsquigarrow$ P-value < 0.005 $\quad$ test not two tail test

$\Rightarrow$ P-value $< \alpha$ $\quad$ So Reject H_0 $(\alpha = 0.01)$

So the proportion are not $\pi_A = 0.40$, $\pi_B = 0.40$, $\pi_C = 0.2$ $(\alpha = 0.01)$.

b.

critical value $\rightsquigarrow$ $\chi^2_\alpha = \chi^2_{0.01} = 9.210$

$\chi^2_\alpha < \chi^2$

$\rightsquigarrow \chi^2 \geq \chi^2_\alpha$

$\Rightarrow$ Reject H_0 $(\alpha = 0.01)$, The proportions are not true π_A ... .

Remark :

To use Goodness of fit test for multinomial populations we assume :

1. The sample taken is random.
2. expected frequencies for all categories ; should satisfy the following :

$e_i \geq 5 \quad \forall i$.

Done 12.1