Lecture 13

Tuesday, February 25, 2020 2:02 PM

- Section 13.5 (r-sample multi-nomial)

Fully Specified / unspecified

- Section 13.6 Full Test of Independence
- Section 13.7 GIF 1-Sample

Pet \ Owner	Female	N-Female	T
Dogs	16	28	44
Cats	24	32	56
Bunny	33	17	50
To	73	77	150

$r = 2,\ c = 3$

		e_{1j}	e_{2j}
$H_0: p_{1\|1} = 0.25$	$p_{1\|2} = 0.25$	18.25	19.25
$p_{2\|1} = 0.25$	$p_{2\|2} = 0.25$	18.25	19.25
$p_{3\|1} = 0.50$	$p_{3\|2} = 0.5$	36.5	38.5

$e_{ij} = n_i \cdot p^o_{j|i}$

- $\chi^2_{OB,1} = \sum_{j=1}^{c} \frac{(o_{1j} - e_{1j})^2}{e_{1j}} = 2.425$
- $\chi^2_{OB,2} = \sum_{j=1}^{c} \frac{(O_{2j} - e_{2j})^2}{e_{2j}} = 24.43$
- $\chi^2_{OB} \approx 26.853$ where $\chi^2_{OB} \sim \chi^2(r(c-1))$

Then since $\chi^2_{0.95}(4) = 9.4$, we reject H_0

Fully-Specified:

- Notation: $p_{j|i}$, $j = \underbrace{1, \ldots, c}_{\text{category}}$, $i = \underbrace{1, \ldots, r}_{\text{\# of MN samples}}$

H_0: $p_{1|1} = p_{1|2} = p_1$

$p_{2|1} = p_{2|2} = p_2$

$p_{3|1} = p_{3|2} = p_3$

where p_1, p_2, p_3 unknown.

$\hat{p}_D = 44/150 \approx 0.293$

$\hat{p}_C = 56/150 \approx 0.373$

$\hat{p}_B = 50/150 \approx 0.33$

Then $e_{ij} = n_i \hat{p}_j$ gives us...

$\hat{e}_{11} = (73)(44/150)$

$\hat{e}_{12} = (73)(56/150)$

$\hat{e}_{11} = (73)(44/150)$

$\hat{e}_{12} = (73)(56/150)$

$\vdots$

$\hat{e}_{23} = (77)(50/150)$

Thus, $X^2_{OB} \sim \chi^2_{((r-1)(c-1))} = 46.103$

$\therefore$ Since $\chi^2_{0.95}(2) = 5.99$, we reject H_0 ✓

Section 13.6:

Recall two events are independent if:

$P(A \cap B) = P(A)P(B)$

$\Leftrightarrow$

$P(A|B) = P(A)$

- $H_0: P_{ij} = P_{i\cdot} \cdot P_{\cdot j}$

	F	NF	$\hat{P}_{i\cdot}$	
D	16/150	28/150	44/150	$\hat{P}_D$
C	24/150	32/150	56/150	$\hat{P}_C$
B	33/150	17/150	50/150	$\hat{P}_B$
$\hat{P}_{\cdot j}$	73/150	77/150		
	$\hat{P}_1$	$\hat{P}_2$		

- Then, the expected counts give...

$\hat{e}_{ij} = N \cdot \hat{P}_{ij}$

$\hat{e}_{i1} = 150 \cdot (44/150)(73/150)$

$\vdots$

under $H_0: P_{ij} = P_{i\cdot} \cdot P_{\cdot j}$

Then $X^2_{OB} = \sum_i \sum_j \frac{(O_{ij} - \hat{e}_{ij})^2}{\hat{e}_{ij}} \sim \chi^2_{(\underbrace{(r-1)}_{=1}\underbrace{(c-1)}_{=2})} \approx 9.44$

$\therefore$ Since $\chi^2_{0.95} = 5.99$, we reject H_0 ✓

Section 13.7 (Goodness of Fit Tests for 1-sample MN):

(A) Completely specified Case

- This is an extension of previous tests to test if the assumption of the distribution under the null is supported by the observed data.

$H_0: X \sim F_X(x)$ where $F_X(x)$ is completely specified (assume values of parameters).

- Suppose sample space has c outcomes $(A_1, A_2, \ldots, A_c)$
- Thus...

$P[X \in A_j] = P_{j0}$ for each A_j

- Thus . . .

$P[X \in A_j] = P_{jo}$ for each A_j

- Given p_{jo}, we can come up with $e_j = P_{jo} \cdot n$ where n = # of total observations

$$\chi^2_{OB} = \sum_{j=1}^{c} \frac{(o_j - e_j)^2}{e_j}$$

$$\sim \chi^2(c-1)$$

Ex M;M Problem

- suppose M;Ms come in 5 colors {R, B, G, O, Y}
- if x_j is counts of M;M's in color j: $P(X=j) = 0.2$
- Suppose the packet contains 100 pieces
  - $X \sim \text{Unif}(1/5)$

  H_0: $X \sim \text{Unif}(1/5)$

	R	B	G	O	Y
o_j	15	31	23	14	17
e_j	20	20	20	20	20

$\Rightarrow \chi^2_{OB} = 18$ $\therefore$ we reject H_0

$\chi^2_{0.95}(4) = 9.48$

Ex 13.7.1

- x is the repair time for some factory component
- H_0: $x \sim \text{Poisson}(3)$
- The total number of components is 40
- The dataset shows us that:

Repair time t (days)	0	1	2	3	4	5	6	> 7
Observed Repair time	1	3	7	6	10	7	6	0
P_{jo}	0.05	0.149	0.224			0.101	0.05	0.034
e_{jo}	(40)(0.05) 5.96		8.96			4.04	2.00	1.36

7.96 7.4