STK 210: Practical 4

1 PROC FREQ: Test of Independence

The FREQ procedure produces one-way and contingency (crosstabulation) tables for categorical variables. PROC FREQ can compute the χ^2-test of **independence for two variables**. The hypotheses for the test are

$$\begin{aligned} H_0 &: \text{ Two variables are independent.} \\ H_1 &: \text{ Two variables are dependent.} \end{aligned}$$

1.1 Example: Beer Preference

Alber's Brewery of Tuscon, Arizona manufactures and distributes three types of beer. A test of independence addresses the question of whether the beer preference (light, regular or dark) is independent of the gender of the beer drinker (male, female). If beer preference depends on the gender of the beer drinker, the firm will tailor its promotions to different target markets. A sample of $n = 150$ beer drinkers is selected and summarised in Table 1.

Table 1: Contingency table of $n = 150$ beer drinkers.

	Beer Preference			
Gender	Light	Regular	Dark	**Total**
Male	20	40	20	**80**
Female	30	30	10	**70**
Total	**50**	**70**	**30**	**150**

1.1.1 Column Percentages

Marginal

- $P\,(\text{Male}) = \dfrac{80}{150} = 0.533\,33$

Partial

- $P\,(\text{Male}|\text{Light}) = \dfrac{20}{50} = 0.4$
- $P\,(\text{Male}|\text{Regular}) = \dfrac{40}{70} = 0.571\,43$
- $P\,(\text{Male}|\text{Dark}) = \dfrac{20}{30} = 0.666\,67$

Note: Males are more inclined to drink darker beer. See column percentages in SAS Output on p.2.

SAS Program:

```
data beer;
input gender $ preference $ count;
datalines;
Male Light     20
Male Regular   40
Male Dark      20
Female Light   30
Female Regular 30
Female Dark    10
;

proc freq data=beer order=data;
tables gender*preference;
weight count;
run;
```

Note: The OPTION `ORDER=DATA` will ensure that the order of the levels of the variables will remain the same as in the input data set.

SAS Output:

```
The FREQ Procedure

Table of gender by preference

gender     preference

Frequency|
Percent  |
Row Pct  |
Col Pct  |Light   |Regular |Dark    |  Total
---------|--------|--------|--------|
Male     |     20 |     40 |     20 |     80
         |  13.33 |  26.67 |  13.33 |  53.33
         |  25.00 |  50.00 |  25.00 |
         |  40.00 |  57.14 |  66.67 |
---------|--------|--------|--------|
Female   |     30 |     30 |     10 |     70
         |  20.00 |  20.00 |   6.67 |  46.67
         |  42.86 |  42.86 |  14.29 |
         |  60.00 |  42.86 |  33.33 |
---------|--------|--------|--------|
Total          50       70       30      150
            33.33    46.67    20.00   100.00
```

1.1.2 Row Percentages

Marginal

- $P\,(\text{Light}) = \dfrac{50}{150} = 0.333\,33$

Partial

- $P\,(\text{Light}|\text{Male}) = \dfrac{20}{80} = 0.25$
- $P\,(\text{Light}|\text{Female}) = \dfrac{30}{70} = 0.428\,57$

Note:

- Probability to drink light beer is not the same for males and females.
- This suggests dependence.
- We would now like to conduct a statistical test, namely Pearson's χ^2- test of independence.

1.1.3 Expected Values

Under the null hypothesis of independence the cell frequencies will marginally reflect the row and column totals i.e. under the null hypothesis of independence

$$\frac{50}{150} = P\,(\text{Light}) = P\,(\text{Light}|\text{Male}) = P\,(\text{Light}|\text{Female}) = \frac{1}{3}$$

It is now fairly straight forward to calculate the expected frequencies under the null hypothesis for:

- $\boxed{\text{GENDER} = \text{Male}}$ and $\boxed{\text{PREFERENCE} = \text{Light}}$

  $$\frac{1}{3}\,(80) = 26.667 = e_{11}$$

  If we compare this with the observed frequency $f_{11} = 20$ we see that there are less males who drink light beer than what is expected under the null hypothesis of independence.

- $\boxed{\text{GENDER} = \text{Female}}$ and $\boxed{\text{PREFERENCE} = \text{Light}}$

  $$\frac{1}{3}\,(70) = 23.333 = e_{21}$$

  If we compare this with the observed frequency $f_{21} = 30$ we see that there are more females who drink light beer than what is expected under the null hypothesis of independence.

Equivalently the expected frequencies under the null hypothesis of independence can be formulated by

$$e_{ij} = \frac{\text{total}\,(\text{row}_i) \times \text{total}\,(\text{col}_j)}{n}$$

therefore

$$e_{11} = \frac{(80)\,(50)}{150} = 26.667 \quad \text{and} \quad e_{21} = \frac{(70)\,(50)}{150} = 23.333$$

Note:

- Under the null hypothesis of independence the observed frequencies (f_{ij}) and expected frequencies (e_{ij}) would be the same.
- The more the observed (f_{ij}) and expected frequencies (e_{ij}) differ from each other, the more reason there is to reject the null hypothesis of independence.

1.1.4 Cell χ^2-values

The deviation between the observed and expected frequency is measured by the cell χ^2-values

$$\text{cell } \chi^2\text{-value} = \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2\,(1)$$

Note: Since $\chi^2_{0.95}\,(1) = 3.841$ (Table III on p.16) we regard a cell χ^2-value of 3 as an indication that the observed and expected frequencies differ from each other.

1.1.5 Test Statistic

The test statistic for independence for a $(r \times c)$ contingency table is

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2\,((r-1)\,(c-1))$$

where

$$\begin{aligned} r &= \text{number of rows} \\ c &= \text{number of columns} \\ f_{ij} &= \text{observed frequency in row } i \text{ and column } j \\ e_{ij} &= \text{expected frequency in row } i \text{ and column } j \end{aligned}$$

Note:

- The larger the value of χ^2 statistic the more reason there is to reject the null hypothesis of independence. (See remark at expected values.)
- In PROC FREQ we can specify the CHISQ option in the TABLES statement so that SAS can perform the test of independence for us.
- The options EXPECTED and CELLCHI2 in the TABLES statement will produce the expected values and cell χ^2-values for us.

SAS Program

```
proc freq data=beer order=data;
tables gender*preference / expected cellchi2 chisq nocum norow nocol nopercent;
weight count;
run;
```

SAS Output

```
The FREQ Procedure

Table of gender by preference

gender          preference

Frequency       |
Expected        |
Cell Chi-Square|Light   |Regular |Dark    |  Total
---------------|--------|--------|--------|
Male           |     20 |     40 |     20 |     80
               | 26.667 | 37.333 |     16 |
               | 1.6667 | 0.1905 |      1 |
---------------|--------|--------|--------|
Female         |     30 |     30 |     10 |     70
               | 23.333 | 32.667 |     14 |
               | 1.9048 | 0.2177 | 1.1429 |
---------------|--------|--------|--------|
Total               50       70       30      150

Statistics for Table of gender by preference

Statistic                     DF       Value      Prob
------------------------------------------------------
Chi-Square                     2      6.1224    0.0468
Likelihood Ratio Chi-Square    2      6.1778    0.0456
Mantel-Haenszel Chi-Square     1      5.8719    0.0154
Phi Coefficient                       0.2020
Contingency Coefficient               0.1980
Cramer's V                            0.2020
```

From the SAS Output:

- We can now read the expected values $e_{11} = 26.667$ and $e_{21} = 23.333$.
- The cell χ^2-value for $\boxed{\text{GENDER} = \text{Male}}$ and $\boxed{\text{PREFERENCE} = \text{Light}}$

$$\frac{(f_{ij} - e_{ij})^2}{e_{ij}} = \frac{(20 - 26.667)^2}{26.667} = 1.6668$$

- Value of the test statistic

$$\begin{aligned}\chi^2 &= \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 1.6667 + 0.1905 + 1 + 1.9048 + 0.2177 + 1.1429 \\ &= 6.1226 \text{ (Not exactly the same as the SAS Output due to rounding errors.)}\end{aligned}$$

From **Output** $\boxed{\chi^2 = 6.1224}$

NB!!! To test the hypotheses we are going to make use of the two approaches that you have done in **STK120** namely

- the **critical value approach** and
- the p-**value approach**.

Critical Value Approach

- Use $\alpha = 0.05$.
- Reject H_0 if $\chi^2 \geq \chi^2_{0.95}(df)$ where $df = (r-1)(c-1) = (1)(2) = 2$ i.e. (Table III on p.16)

$$\boxed{\chi^2_{0.95}(2) = 5.991}$$

- **SAS function:**

`quantile('chisquare',0.95,2)` `5.9914645`

- Since $(\chi^2 = 6.1224) \geq (\chi^2_{0.95}(2) = 5.991)$ we reject H_0 at 5% level of significance.
- **Conclusion:** GENDER and beer PREFERENCE are dependent.

p**-value Approach**

- Use $\alpha = 0.05$.
- Reject H_0 if p-value ≤ 0.05. From the SAS Output

$$\boxed{p\text{-value} = 0.0468}$$

- **SAS function:**

`1-cdf('chisquare',6.1224,2)` `0.0468315`

- Since (p-value $= 0.0468) \leq 0.05$ we reject H_0 at 5% level of significance.
- **Conclusion:** GENDER and beer PREFERENCE are dependent.

1.1.6 Data

PROC FREQ can use **raw data** or **cell count data**. Up to now we have made use of cell count data, because we used the frequency data that was summarised in the Table 1. Usually we get data in a **raw format**. We will then need PROC FREQ to count the data for us. See SAS Program and SAS Output below.

SAS Program:

```
data beer;
input individual gender $ preference $;
datalines;
1   Female  Light
2   Female  Regular
3   Male    Regular
4   Female  Light
5   Female  Light
.
147 Male    Light
148 Male    Regular
149 Female  Light
150 Female  Regular
;

proc freq data=beer;
tables gender*preference / chisq expected cellchi2;
run;
```

SAS Output:

```
The FREQ Procedure

Table of gender by preference

gender          preference

Frequency      |
Expected       |
Cell Chi-Square|
Percent        |
Row Pct        |
Col Pct        |Dark    |Light   |Regular |  Total
---------------|--------|--------|--------|
Female         |     10 |     30 |     30 |     70
               |     14 | 23.333 | 32.667 |
               | 1.1429 | 1.9048 | 0.2177 |
               |   6.67 |  20.00 |  20.00 |  46.67
               |  14.29 |  42.86 |  42.86 |
               |  33.33 |  60.00 |  42.86 |
---------------|--------|--------|--------|
Male           |     20 |     20 |     40 |     80
               |     16 | 26.667 | 37.333 |
               |      1 | 1.6667 | 0.1905 |
               |  13.33 |  13.33 |  26.67 |  53.33
               |  25.00 |  25.00 |  50.00 |
               |  66.67 |  40.00 |  57.14 |
---------------|--------|--------|--------|
Total                30       50       70      150
                  20.00    33.33    46.67   100.00
```

```
Statistics for Table of gender by preference

Statistic                     DF       Value      Prob
------------------------------------------------------
Chi-Square                     2      6.1224    0.0468
Likelihood Ratio Chi-Square    2      6.1778    0.0456
Mantel-Haenszel Chi-Square     1      0.0794    0.7781
Phi Coefficient                       0.2020
Contingency Coefficient               0.1980
Cramer's V                            0.2020

Sample Size = 150
```

Note:

- The levels of the two categorical variables are listed alphabetically.
- Make sure that you understand all the values that are cross classified in the cells.

2 PROC IML: The SAMPLE function

The SAMPLE function generates a random sample of the elements of the vector `x`. The function can sample from `x` with replacement or without replacement. The function can sample from `x` with equal probability or with unequal probability.

Syntax:

`SAMPLE(x<,n><,method><,prob>);`

`x` is a matrix that specifies the sample space i.e. the sample is drawn from the elements of x.

`n` specifies the number of times to sample. The argument can be a scalar or a two-element vector.

- If this argument is omitted, then the number of elements of `x` is used.
- If `n` is a scalar, then it represents the sample size, which is the number of independent draws from the population. This value determines the **number of columns** in the output matrix.
- If `n` is a two-element vector, the **first element** represents the **sample size**. The **second element** specifies the **number of samples**, which is the number of rows in the output matrix. If the sampling is without replacement, then `n[1]` must be less than or equal to the number of elements in `x`.

`method` is an optional argument that specifies how sampling is performed. The following are valid options:

- "`Replace`" specifies simple random sampling with replacement. This is the default value.
- "`NoReplace`" specifies simple random sampling without replacement. The elements in the samples might appear in the same order as in `x`.
- "WOR" specifies simple random sampling without replacement. After elements are randomly selected, their order is randomly permuted.

`prob` is a vector with the same number of elements as `x`. The vector specifies the sampling probability for the elements of `x`. The SAMPLE function internally scales the elements of `prob` so that they sum to unity.

Note:

- The SAMPLE function uses the random seed that is set by the RANDSEED function.
- The prob argument specifies the probabilities that are used when sampling from `x`. When method is "Replace," the probabilities do not change during the sampling. However, when method is "NoReplace," the probabilities are renormalized after each selection.

SAS Program

```
proc iml;
x = 1:5;
print x;
call randseed(111,1); s1=sample(x);
call randseed(222,1); s2=sample(x,5,"Replace",{0.6 0.1 0.0 0.1 0.2});
call randseed(333,1); s3=sample(x,3,"NoReplace");
call randseed(444,1); s4=sample(x,{3,10},"Replace");
call randseed(555,1); s5=sample(x,{3,10},"NoReplace");
print s1,s2,s3,'Replacement' s4 'No replacement' s5;
```

SAS Output

```
                       x
      1         2         3         4         5

                       s1
      2         3         5         1         4

                       s2
      1         2         1         1         1

             s3
      1         2         5

                 s4                                         s5
Replacement       2         4         4 No replacement       1         2         5
                  4         1         4                      1         5         4
                  2         4         1                      5         2         3
                  5         5         1                      1         4         3
                  2         5         3                      5         2         3
                  3         5         4                      4         2         3
                  1         3         5                      4         2         3
                  3         4         5                      4         2         5
                  2         5         4                      1         4         3
                  2         4         2                      5         4         3
```

3 Exercise

1. Consider an experiment that consists of recording the birthday for each of $n = 20$ randomly selected persons.

   Assume: There are no leap years and all birthdays are equally likely.

   Define the events:

$$\begin{aligned} A &= \text{each person has a different birthday.} \\ B &= \text{at least two people share the same birthday.} \end{aligned}$$

   (a) Find the number of sample points in the sample space S.

$$(365)^{20} = 1.7614 \times 10^{51}$$

   (b) Find the number of sample points in A.

$$P_{20}^{365} = 1.0367 \times 10^{51}$$

   (c) If we assume that all birthdays are equally likely, what is the probability that each person in the $n = 20$ sample has a **different** birthday?

$$P(A) = \frac{P_{20}^{365}}{(365)^{20}} = 0.58856$$

   (d) Calculate the probability that at least two people share the same birthday. (Complement rule.)

$$P(B) = 1 - P(A) = 1 - 0.58856 = 0.41144$$

   (e) Use a DO LOOP to calculate the probability to find at least two people with the same birthday for the following sample sizes: **Complete by using SAS!**

n	$P(A) = \frac{P_n^{365}}{(365)^n}$	$P(B) = 1 - P(A)$
10	0.8830518	
15		0.2529013
20		
25	0.4313003	0.5686997
30		
35		0.8143832
40		
45		
50	0.0296264	
55		
60		
65		0.9976831
70		

   You can use the following in PROC IML:

```
do n=10 to 70 by 5;
     statements
end;
```

2. Two socks are selected at random from a drawer containing five brown socks and three green socks.

Define the events:

$$B_1 = \text{First sock is a brown sock} \quad \text{and} \quad B_2 = \text{Second sock is a brown sock}$$

and

$$G_1 = \text{First sock is a green sock} \quad \text{and} \quad G_2 = \text{Second sock is a green sock}$$

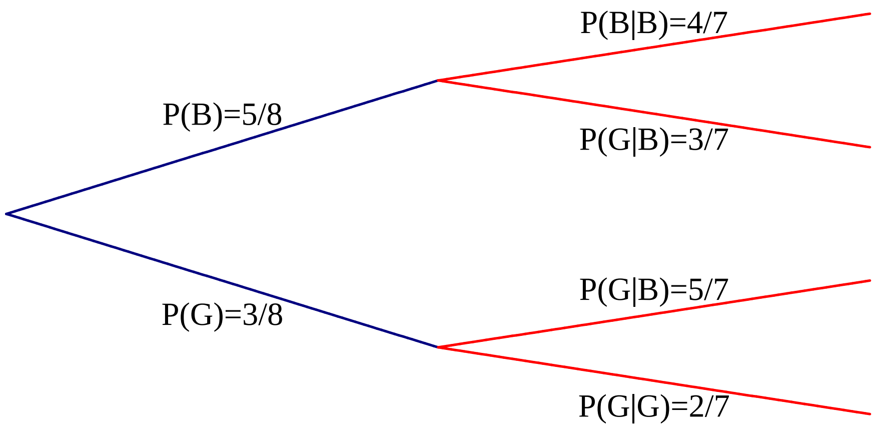
Calculate the probabilities:

$$\begin{aligned} P(B_1B_2) &= P(BB) = \\ P(B_1G_2) &= P(BG) = \\ P(G_1B_2) &= P(GB) = \\ P(G_1G_2) &= P(GG) = \end{aligned}$$

for the four scenarios listed below, i.e. 2(a) i & ii and 2(b) i & ii.

(a) Suppose the two socks are removed in succession i.e. **without replacement**.

i. Determine the **theoretical distribution**. Use probability theory to obtain the theoretical probabilities.



ii. Determine the **empirical distribution**.

- Use the SAMPLE function in PROC IML with the RANDSEED call with a seed of $\boxed{612}$ and let

$$S = \{B, B, B, B, B, G, G, G\}$$

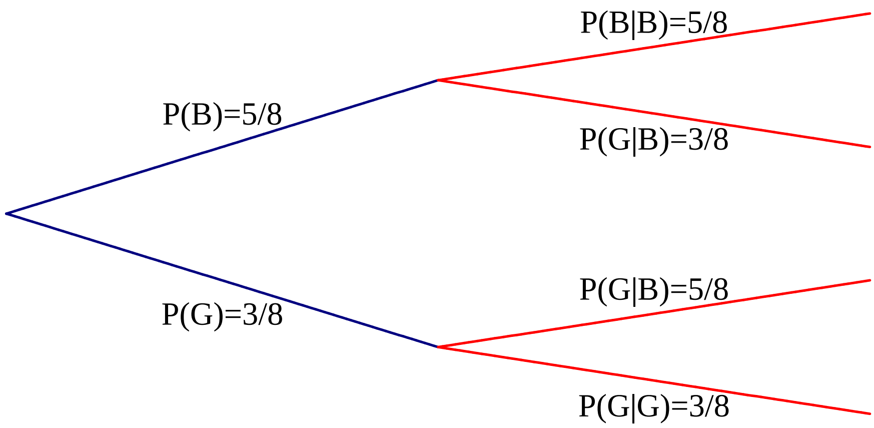
denote the sample space.

- Generate the matrix $\mathbf{X}$ with 10000 samples and a sample size of 2. For every row of the matrix $\mathbf{X}$ two socks are selected without replacement. Print row numbers 101 to 105 of the matrix $\mathbf{X}$.
- Create the SAS data set SOCKS with $n = 10000$ observations and variables SOCK1 and SOCK2. SOCK1 is the first sock selected from the drawer and SOCK2 is the second.
- Use PROC FREQ to obtain the empirical probabilities.

Note: Two events are **dependent**.

(b) Suppose the two socks are removed **with replacement**.

i. Determine the **theoretical distribution**. Use probability theory to obtain the theoretical probabilities.



ii. Determine the **empirical distribution**.

- Use the SAMPLE function in PROC IML with the RANDSEED call with a seed of $\boxed{238}$ and let

$$S = \{B, B, B, B, B, G, G, G\}$$

denote the sample space.

- Generate the matrix $\mathbf{X}$ with 10000 samples and a sample size of 2. For every row of the matrix $\mathbf{X}$ two socks are selected with replacement. Print the first 5 elements of the matrix $\mathbf{X}$.
- Create the SAS data set SOCKS with $n = 10000$ observations and variables SOCK1 and SOCK2. SOCK1 is the first sock selected from the drawer and SOCK2 is the second.
- Use PROC FREQ to obtain the empirical probabilities.

Note: Two events are **independent**.

Solution: $\boxed{\text{Do you agree with the following answers?}}$

(a) **Without replacement:**

Probability	Theoretical	Empirical
$P(BB)$	0.3571	0.3646
$P(BG)$	0.2679	0.2677
$P(GB)$	0.2679	0.2618
$P(GG)$	0.1071	0.1059

(b) **With replacement:**

Probability	Theoretical	Empirical
$P(BB)$	0.3906	0.3942
$P(BG)$	0.2344	0.2380
$P(GB)$	0.2344	0.2312
$P(GG)$	0.1406	0.1366

3. Due to rising health insurance costs, 43 million people in the United States go without health insurance (Time, December 1, 2003). Sample data representative of the national health insurance coverage are shown in the following contingency table.

Health Insurance	**Age** 18-34 G	35+ $\overline{G}$
Yes H	60	105
No $\overline{H}$	20	15

Define the events:

$$H = \text{Person has health insurance}$$
$$G = \text{Person is in younger age category, i.e. 18-34}$$

(a) Create the SAS data set $\boxed{\text{HEALTH}}$ with variables INSURANCE and AGE.

(b) Use PROC FREQ to answer the following questions.

i. Use one-way frequency distributions to detemine

A. $P(H)$ — 0.825

B. $P(G)$ — 0.4

ii. Use two-way frequency distribution to determine

A. $P(H|G)$ — 0.75

B. $P\left(H|\overline{G}\right)$ — 0.875

C. Are the events H and G independent?

- No, older people are more inclined to have health insurance.
- $P(H) \neq P(H|G) \neq P\left(H|\overline{G}\right)$

iii. Use two-way frequency distribution to determine

A. $P\left(\overline{G}|H\right)$ — 0.6364

B. $P\left(\overline{G}|\overline{H}\right)$ — 0.4286

C. Are the events H and G independent?

- No, prob to be "old" is higher when you have health insurance.
- $P\left(\overline{G}\right) \neq P\left(\overline{G}|H\right) \neq P\left(\overline{G}|H\right)$

iv. Use two-way frequency distribution to determine

A. $P(H \cap G)$ — 0.3

B. $P\left(H \cap \overline{G}\right)$ — 0.525

C. $P\left(\overline{H} \cap G\right)$ — 0.1

D. $P\left(\overline{H} \cap \overline{G}\right)$ — 0.075

v. Check the equalities:

A. $P(H \cap G) = P(H)P(G)$ — $P(H)P(G) = (0.825)(0.4) = 0.33 \neq 0.3$

B. $P\left(H \cap \overline{G}\right) = P(H)P\left(\overline{G}\right)$ — $P(H)P\left(\overline{G}\right) = 0.825(0.6) = 0.495 \neq 0.525$

C. Are the events H and G independent?

- No, joint probabilities are not the same as product of marginal probabilities.

vi. Consider the hypotheses

H_0 : Possession of health insurance is **independent** of age.

H_1 : Possession of health insurance is **dependent** of age.

Let: $\alpha = 0.05$

A. Give the observed frequency for a person that is at least 35 years to have health insurance i.e. f_{12}. 105

B. Give the expected frequency for a person that is at least 35 years to have health insurance i.e. e_{12}. 99

C. Are the events H and A independent if you compare f_{12} and e_{12}?

- No, $f_{12} \neq e_{12}$. There are more people in older category with health insurance $f_{12} = 105$ than what is expected under independence $e_{12} = 99$.

D. Give the cell χ^2 value for the cell in the second row first column. Are the events H and A independent? 2.5714

- No, under independence the cell χ^2 value would have been zero. Some indication of dependence.

E. Give the value of the test statistic. $\chi^2 = 5.1948$

F. Use the p-value approach to draw a conclusion. (Give the p-value.)

- $(p\text{-value} = 0.0227) < 0.05$
- Reject H_0 at 5% level of significance.
- Two variables are dependent.

G. Use the critical value approach to draw a conclusion. (Give the critical value. Use χ^2-table on p.16.)

- $(\chi^2 = 5.1948) \geq \chi^2_{0.95}(1) = 3.841$
- Reject H_0 at 5% level of significance.
- Two variables are dependent.

(c) Use PROC IML to create the matrix

$$\mathbf{F} = \begin{pmatrix} 60 & 105 \\ 20 & 15 \end{pmatrix}$$

from the SAS data set HEALTH. Print the matrix $\mathbf{F}$.

Note: You can check all your answers for this question with the Output of PROC FREQ.

i. Use PROC IML to create the matrix $\mathbf{P}$ with the 4 probabilities listed below:

A. $P(H \cap G)$

B. $P\left(H \cap \overline{G}\right)$

C. $P\left(\overline{H} \cap G\right)$

D. $P\left(\overline{H} \cap \overline{G}\right)$

ii. Use PROC IML to create the column vector $\mathbf{H}$ with the 2 probabilities listed below.

A. $P(H)$

B. $P(\overline{H})$

iii. Use PROC IML to create the row vector $\mathbf{G}$ with the 2 probabilities listed below.

A. $P(G)$

B. $P(\overline{G})$

iv. Use PROC IML to calculate the conditional probabilities:

A. $P(H|G)$

B. $P(H|\overline{G})$

v. Use PROC IML to calculate the conditional probabilities:

A. $P(\overline{G}|H)$

B. $P(\overline{G}|\overline{H})$

vi. Consider the hypothesis

H_0 : Whether a person has health insurance is **independent** of age

H_1 : Whether a person has health insurance is **dependent** on age

Let: $\alpha = 0.05$

A. Calculate the matrix

$$\mathbf{P}_0 = \mathbf{HG}$$

with the **expected probabilities** under the null hypothesis of independence.

Note: Under independence

- $P(H \cap G) = P(H)\,P(G)$
- $P(H \cap \overline{G}) = P(H)\,P(\overline{G})$
- $P(\overline{H} \cap G) = P(\overline{H})\,P(G)$
- $P(\overline{H} \cap \overline{G}) = P(\overline{H})\,P(\overline{G})$

B. Calculate the matrix

$$\mathbf{F}_0 = n\mathbf{P}_0$$

with the **expected frequencies** under the null hypothesis of independence.

C. Calculate the matrix $\mathbf{X}$ with the cell χ^2 values.

D. Calculate the value of the test statistic.

E. Calculate the value of the critical value.

F. Calculate the value of the p-value.

Table III: Percentiles of χ^2 distribution, values of $\chi^2_\gamma(\nu)$ where $\gamma = \int_0^{\chi^2_\gamma(\nu)} f(x;\nu)\,dx.$

	γ							
ν	**0.005**	**0.01**	**0.025**	**0.05**	**0.95**	**0.975**	**0.99**	**0.995**
1	0.0000393	0.000157	0.000982	0.00393	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672