**Your Name: Hendi Kushta**

**IST772 Week 2 Class Exercise – Practice with Conditional Probabilities**

**Instructions:** **Post this document with your code, results, and answers to all questions in Blackboard.**

The table below shows accident reports from three different factories over the past month. Four types of accidents are represented. Each cell contains a count of the number of accidents of the given type at the particular factory:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Accidents** | *Factory 1* | *Factory 2* | *Factory 3* |  |
| *Vehicle* | 0 | 6 | 4 | 10 |
| *Spill* | 6 | 0 | 6 | 12 |
| *Equipment* | 6 | 4 | 5 | 15 |
| *Injury* | 4 | 9 | 0 | 13 |
|  | 16 | 19 | 15 | 50 |

1. Add marginal totals to the table above for cross checking with your R results. You can just type those above.

1. **Recreate the matrix in R using the following code**:

accMatrix <- matrix(data=c(0,6,4,6,0,6,6,4,5,4,9,0),nrow=4,byrow=T, dimnames=list(c("Vehicle","Spill","Equipment","Injury"), c("Factory 1", "Factory 2", "Factory 3")))

A picture containing chart

Description automatically generated

1. **Create a copy of the matrix that contains proportions instead of counts.** One helpful function that can be called on the whole matrix is sum( ).

Table

Description automatically generated

Proportions matrix can be considered as the probability matrix which is calculated by dividing each entry by the total of the entries in the table.

1. **Calculate marginal totals for the proportion matrix.** Two helpful functions that can be used are rowSums( ) and colSums( ).

A picture containing text

Description automatically generated

Text

Description automatically generated

Table

Description automatically generated

1. OSHA is auditing the factory that has the worst accident record overall.

a. **Add a comment in your code indicating which factory that is.**

b. **For that factory, list the proportions from #3 of each type of accident,** using the [ ] subsetting technique.For example, you could show the first column of a matrix named Matrix with this command: Matrix[ ,1]

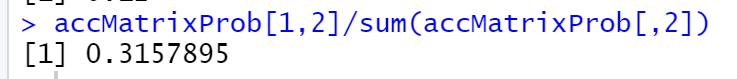
Ans.

1. Since factory 2 has the highest sum of margins of 0.38 in terms of accident proportions, it has the worst accident record overall.
2. The proportions for factory 2:

Text

Description automatically generated with low confidence

1. Putting your focus solely on ***accidents at that factory*, what's the probability of vehicle accidents at that factory?** Write a line of R code that displays the result and include a comment with your answer.



The probability of the occurrence of accidents because of vehicles in factory 2 is 31.6%.

1. The insurance company for these factories wants to understand the most prevalent type of accident across all factories.

a. **Add a comment in your code indicating which type of accident that is.**

b. **For that type of accident, list the proportions from #3 for each factory,** using the [ ] subsetting technique.

Ans.

Equipment based accident is the most prevalent type of accident. This is because it has the highest marginal total across all the factories.

a.

A picture containing text

Description automatically generated

b.

Text

Description automatically generated

1. Putting your focus solely on ***that kind of accident*, what's the probability of that kind of accident at each factory?** Write a line of R code that displays the result and include a comment with your answer.

Ans.

For the probability of occurrence of equipment type of accident at each factory, the probability of occurrence of this kind of accident is divided by the sum of the same for each factory.

Text

Description automatically generated with medium confidence