HW8-Q3

BNarayanan

2024-12-02

```r
dat <- expand.grid(factory = c("East", "West"), accident = c("No", "Yes"))
dat$y <- c(645, 1275, 28, 31)

tab <- matrix(dat$y, nrow = 2,
              dimnames = list(factory = c("East", "West"),
                              accident = c("No", "Yes")))
```

Part A

```r
n_total <- sum(tab)

n_accidents <- sum(tab[, "Yes"])
n_west <- sum(tab["West", ])

p_accident <- n_accidents / n_total
p_west <- n_west / n_total

# Expected accidents in the West factory
expected_west_accidents <- p_accident * p_west * n_total
expected_west_accidents
```

```
## [1] 38.93583
```

Part B

```r
expected_counts <- chisq.test(tab)$expected
expected_counts
```

```
##        accident
## factory        No       Yes
##    East   652.9358 20.06417
##    West  1267.0642 38.93583
```

Part C

```r
pi_factory <- rowSums(tab) / n_total
pi_accident <- colSums(tab) / n_total

# General formula for expected counts
expected_mij <- outer(pi_factory, pi_accident) * n_total
expected_mij
```

```
##               No       Yes
## East   652.9358 20.06417
## West  1267.0642 38.93583
```

Part D

```
log_mij <- log(n_total) + log(pi_factory) %o% log(pi_accident)
log_mij
```

```
##              No       Yes
## East 7.622992 11.37927
## West 7.602926  9.05035
```

Part E

```
log_pi_ij <- log(pi_factory) %o% log(pi_accident)
log_pi_ij
```

```
##                No       Yes
## East 0.03264547 3.788922
## West 0.01257943 1.460003
```