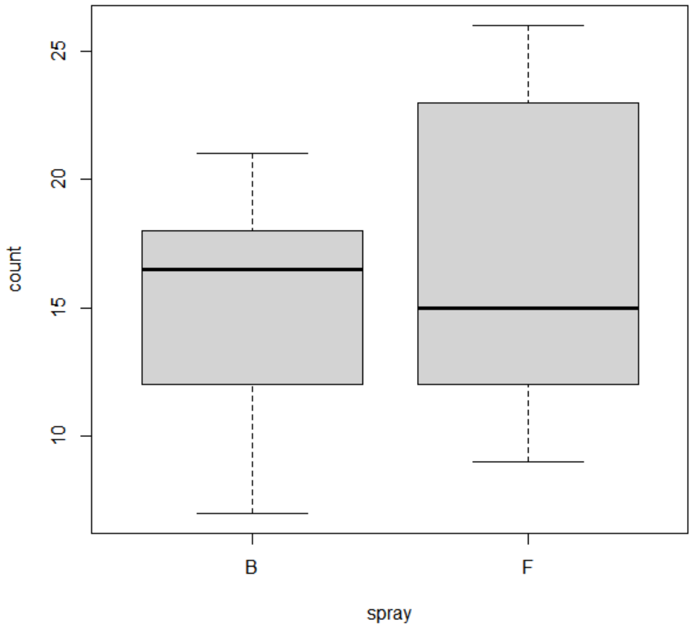
8장 보고서

R의 InsectSprays에서 B,F를 뿌릴 때, 죽는 벌레 수가 동일한지 검정하기 위하여, 유의수준 0.05에서 이표본 T-검정을 실시해보 자. 그림1은 자료의 상자도표이다.



두 스프레이를 뿌릴 때 죽은 평균 벌레수가 동일한지 알아보기 위하여, 다음과 같이 가설을 세우자.

$$H_0 : \mu_B = \mu_F \quad H_1 : \mu_B \neq \mu_F$$

표본크기는 각각 $n_1 = 12$, $n_2 = 12$ 이고, 표본평균은 $\bar{x} = 15.3333$, $\bar{y} = 16.6667$ 이고, 표본표준편차는 $S_x = 4.271115$, $S_x = 6.213378$ 이다. 등분산검정에 대한 유의확률 p=0.2294 가 유의수준 α = 0.05보다 크므로, 등분산이다 등분산 T-검정을 이용하여 계산한 평균차이 $(\mu_B - \mu_F)$ 에 대한 95% 신뢰구간은 (-5.847224, 3.180557) 이고, 검정통계량은 T=-0.61259 이며, 유의확률은 p =0.5464 이다. 따라서 유의수준 0.05 에서 귀무가설을 기각하지 않는다. 즉, 유의수준 0.05에서 살충제 B와 F의 효과는 같다.

```
# R 코드
InsectSprays
boxplot(count~spray, data = InsectSprays)
xy <- subset(InsectSprays, spray=="B" | spray=="F")
xy$spray <- droplevels(xy$spray)
boxplot(count~spray, data=xy)

x <- subset(xy, spray=="B", c(count))$count
y <- subset(xy, spray=="F", c(count))$count
```

```
var.test(x,y)
t.test(x, y, var.equal = T)
t.test(x,y)
mean(x)
mean(y)
sd(x)
sd(y)
length(x)
length(y)
```

```
# 결과
InsectSprays
    count spray
1       10      A
2        7      A
3       20      A
4       14      A
5       14      A
6       12      A
7       10      A
8       23      A
9       17      A
10      20      A
11      14      A
12      13      A
13      11      B
14      17      B
15      21      B
16      11      B
17      16      B
18      14      B
19      17      B
20      17      B
21      19      B
```

22	21	B
23	7	B
24	13	B
25	0	C
26	1	C
27	7	C
28	2	C
29	3	C
30	1	C
31	2	C
32	1	C
33	3	C
34	0	C
35	1	C
36	4	C
37	3	D
38	5	D
39	12	D
40	6	D
41	4	D
42	3	D
43	5	D
44	5	D
45	5	D
46	5	D
47	2	D
48	4	D
49	3	E
50	5	E
51	3	E
52	5	E
53	3	E
54	6	E
55	1	E
56	1	E
57	3	E
58	2	E
59	6	E
60	4	E

```
61      11        F
62       9        F
63      15        F
64      22        F
65      15        F
66      16        F
67      13        F
68      10        F
69      26        F
70      26        F
71      24        F
72      13        F
> boxplot(count~spray, data = InsectSprays)
> xy <- subset(InsectSprays, spray=="B" | spray=="F")
> xy$spray <- droplevels(xy$spray)
> boxplot(count~spray, data=xy)
>
> x <- subset(xy, spray=="B", c(count))$count
> y <- subset(xy, spray=="F", c(count))$count
>
> var.test(x,y)

        F test to compare two variances

data:  x and y
F = 0.47253, num df = 11, denom df = 11, p-value = 0.2294
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1360301 1.6414182
sample estimates:
ratio of variances
         0.4725275

> t.test(x, y, var.equal = T)

        Two Sample t-test

data:  x and y
t = -0.61259, df = 22, p-value = 0.5464
```

```
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -5.847224  3.180557
sample estimates:
mean of x mean of y
 15.33333  16.66667

> t.test(x,y)

	Welch Two Sample t-test

data:  x and y
t = -0.61259, df = 19.498, p-value = 0.5472
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -5.881042  3.214376
sample estimates:
mean of x mean of y
 15.33333  16.66667

> mean(x)
[1] 15.33333
> mean(y)
[1] 16.66667
> sd(x)
[1] 4.271115
> sd(y)
[1] 6.213378
> length(x)
[1] 12
> length(y)
[1] 12
>
```