

練習1

(1)

```
> (3^2-5)/4
```

```
[1] 1
```

(2)

```
> a <- 2
```

```
> b <- 3
```

```
> c <- 1
```

```
> (-b+(b^2-4*a*c)^(-2))/(2*a)
```

```
[1] -0.5
```

(3)

```
> d <- c(4, 9, 25, 36)
```

```
> sqrt(d)
```

```
[1] 2 3 5 6
```

練習2

(1)

```
> dat <- read.csv("grade1.csv")
```

```
> apply(dat[,2:5], 2, sd)
```

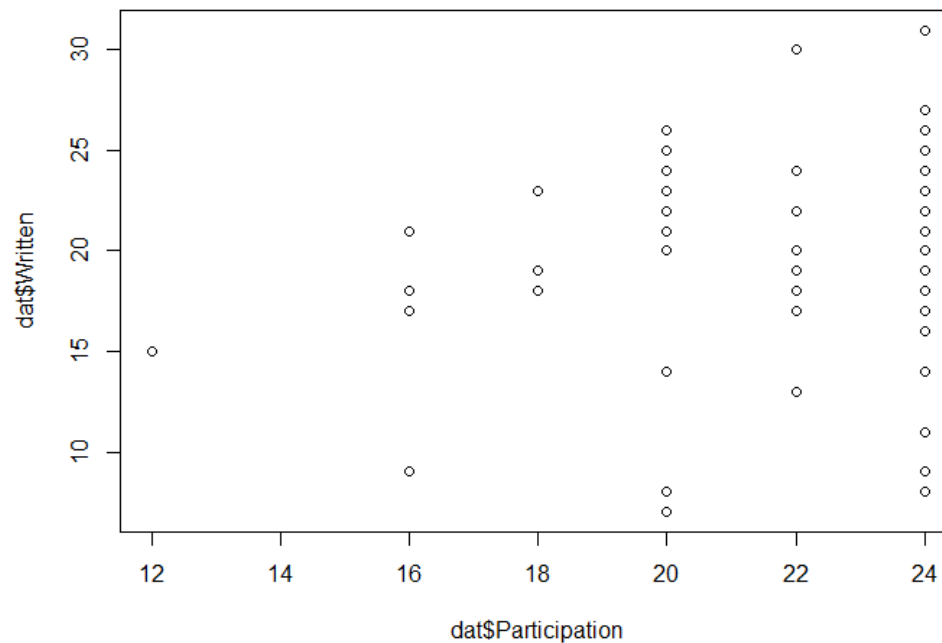
Report	Participation	Written	Practice
4.875485	2.599190	5.414390	3.700427

(2)

```
> plot(dat[,3], dat[,4])
```

←この2つは同じ。列番号より、補完機能が使え直接指定（\$を使う）の方が楽か。

```
> plot(dat$Participation, dat$Written)
```



標準偏差から、平常点は比較的ばらつきは少なく、筆記テストでそれが大きくなっていることがわかる。実技試験も10点満点であることを考えると、ややばらつきが大きいように見える。

散布図からは、平常点にかかわらず、筆記試験の点は大きくばらついていることがわかる。出席しているだけでは、実力がつかないということである。

練習3

(1)

```
> mcn <- read.csv("machine5.csv")
```

```
> tapply(mcn$Weight, list(mcn$MachineID, mcn$Period), mean)
```

	Day	Night
A	100.0063	100.1102
B	99.6450	99.3340

```
> tapply(mcn$Weight, list(mcn$MachineID, mcn$Period), var)
```

	Day	Night
A	0.01058972	0.02407285
B	0.04744915	0.09769898

```
> tapply(mcn$Weight, list(mcn$MachineID, mcn$Period), sd)
```

	Day	Night
A	0.1029064	0.1551543
B	0.2178283	0.3125684

```
> boxplot(mcn$Weight~mcn$MachineID)
```

```
> stripchart(mcn$Weight~mcn$MachineID, vertical = TRUE, pch = 21, col = "maroon", bg = "orange", method = "jitter", add = TRUE)
```

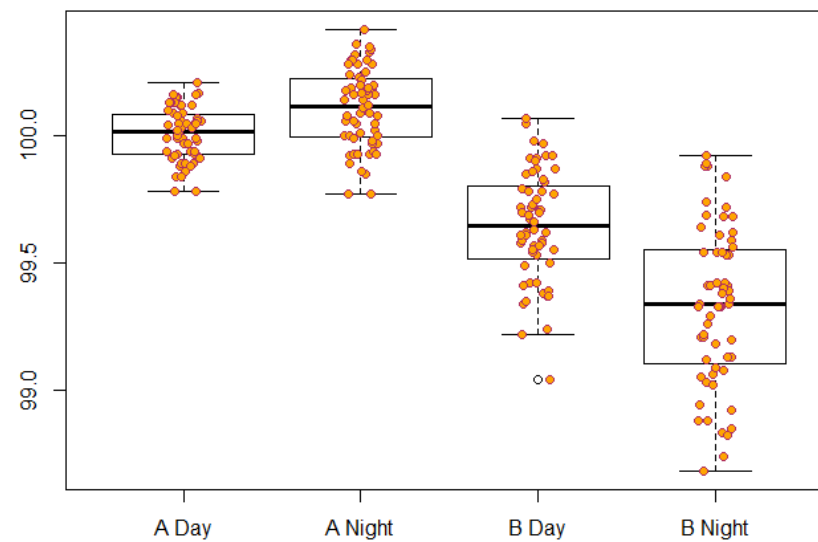
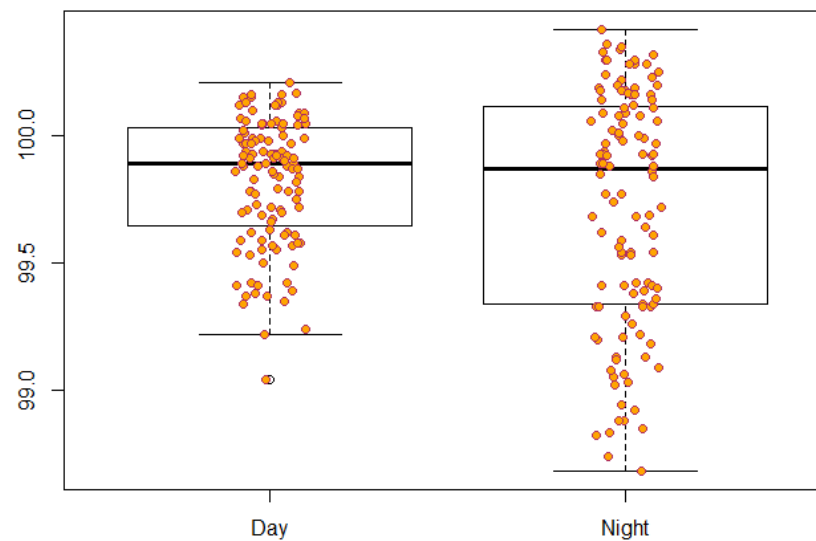
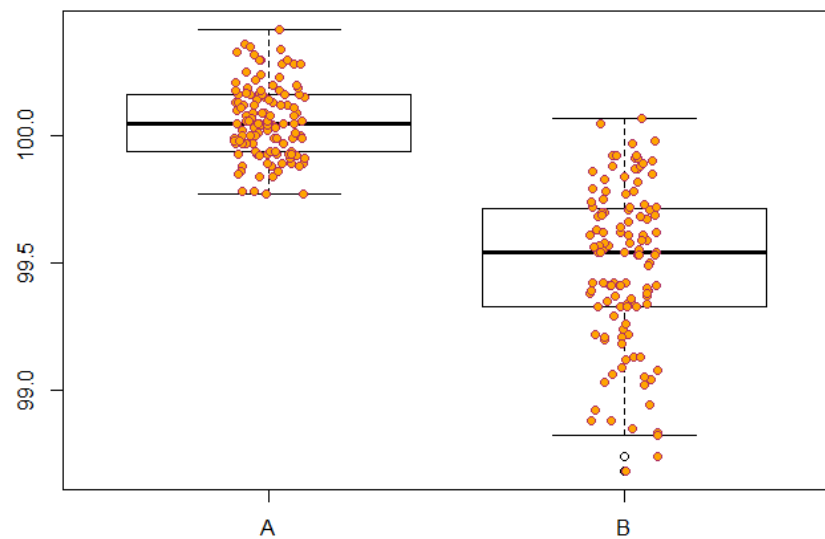
```
> boxplot(mcn$Weight~mcn$Period)
```

```
> stripchart(mcn$Weight~mcn$Period, vertical = TRUE, pch = 21, col = "maroon", bg = "orange", method = "jitter", add = TRUE)
```

```
> case <- paste(mcn$MachineID, mcn$Period)
```

```
> boxplot(mcn$Weight~case)
```

```
> stripchart(mcn$Weight~case, vertical = TRUE, pch = 21, col = "maroon", bg = "orange", method = "jitter", add = TRUE)
```



クロス表の数値をみると、BはAより不安定で、昼と夜を比べると夜の方が悪い。

4つに場合分けすると、箱ひげ図からも、Bの夜がもっとも全体にばらつきが大きく、製品も軽くなっていることがわかる。

(2)

```
> cst <- read.csv("customer.csv")  
> table(cst$Branch, cst$Sex)
```

	女	男
岡山	16	14
岡山南	12	18
玉島	9	21
児島	21	9
倉敷	14	16

岡山支店と倉敷支店の来客は、男女ほぼ同じぐらいであるが、岡山南支店と玉島支店は男性が多く、児島支店では女性が多いことがわかる。

練習4

```
> str <- read.csv("store.csv", row.names="branch")
> lm.str <- lm(sales~., data=str)
> summary(lm.str)
```

Call:

```
lm(formula = sales ~ ., data = str)
```

Residuals:

Min	1Q	Median	3Q	Max
-20.222	-8.507	-1.176	5.645	34.074

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36.87296	32.99290	1.118	0.2962
numpass	0.03676	0.01343	2.738	0.0255 *
minutes	-3.34342	1.27001	-2.633	0.0301 *
area	-0.29716	0.67257	-0.442	0.6703
parkcar	2.44795	1.00274	2.441	0.0405 *
numwork	2.75809	2.21669	1.244	0.2486
kinds	0.03024	0.23975	0.126	0.9027

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 18.74 on 8 degrees of freedom

Multiple R-squared: 0.9549, Adjusted R-squared: 0.9211
F-statistic: 28.24 on 6 and 8 DF, p-value: 5.762e-05

```
> x1 <- c(1, 1956, 3, 88, 42, 10, 120)
> x2 <- c(1, 1300, 12, 90, 45, 10, 100)
> x3 <- c(1, 1423, 8, 42, 36, 10, 90)
> t(lm.str$coef) %*% x1
      [,1]
[1,] 206.6213
> t(lm.str$coef) %*% x2
      [,1]
[1,] 158.5597
> t(lm.str$coef) %*% x3
      [,1]
[1,] 168.385
```

回帰式は, $y = 36.87296 + 0.03676x_1 - 3.34342x_2 - 0.29716x_3 + 2.44795x_4 + 2.75809x_5 + 0.03024x_6$ となる。

これより, 各候補地の売上予測は, 206.6 万円, 158.6 万円, 168.4 万円となるので, 候補 1 を第 1 候補とするのがよい。

練習5

```
> epw <- read.csv("EPower1.csv")
> cor(epw$Temp, epw$Power.Max)
[1] 0.4973176
> lm.epw <- lm(epw$Power.Max~epw$Temp)
> summary(lm.epw)
```

Call:

```
lm(formula = epw$Power.Max ~ epw$Temp)
```

Residuals:

Min	1Q	Median	3Q	Max
-198.20	-61.80	20.65	74.23	120.56

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	139.588	173.589	0.804	0.424
epw\$Temp	23.649	5.326	4.440	3.92e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 91.24 on 60 degrees of freedom

Multiple R-squared: 0.2473, Adjusted R-squared: 0.2348

F-statistic: 19.72 on 1 and 60 DF, p-value: 3.921e-05


```

> epw[45:47,3] <- epw[1,3]          ←お盆で休みとなるのは, 45~47 番目。その値を H (要素(1,3)の値) で置き換える。
> ep.n <- epw[epw$Day==" N",]        ←N である要素を取り出し, 新しいオブジェクト ep.n をつくる。
> ep.h <- epw[epw$Day==" H",]        ←同様に, H であるものを取り出し, 新しいオブジェクト ep.h をつくる。
> lm.ep.n <- lm(ep.n$Power.Max~ep.n$Temp)
> summary(lm.ep.n)

```

Call:

```
lm(formula = ep.n$Power.Max ~ ep.n$Temp)
```

Residuals:

Min	1Q	Median	3Q	Max
-85.176	-20.342	6.394	22.845	60.373

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	252.37	84.53	2.986	0.005 **
ep.n\$Temp	22.04	2.59	8.512	3.06e-10 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 36.91 on 37 degrees of freedom

Multiple R-squared: 0.662, Adjusted R-squared: 0.6528

F-statistic: 72.45 on 1 and 37 DF, p-value: 3.055e-10

```
> lm.ep.h <- lm(ep.h$Power.Max~ep.h$Temp)
> summary(lm.ep.h)
```

Call:

```
lm(formula = ep.h$Power.Max ~ ep.h$Temp)
```

Residuals:

Min	1Q	Median	3Q	Max
-96.429	-45.889	8.579	46.008	71.711

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	16.046	189.396	0.085	0.933285
ep.h\$Temp	24.297	5.826	4.170	0.000432 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 55.3 on 21 degrees of freedom

Multiple R-squared: 0.453, Adjusted R-squared: 0.427

F-statistic: 17.39 on 1 and 21 DF, p-value: 0.0004323

回帰式は、それぞれ、 $y=252.37+22.04x_1$, $y=16.05+24.306x_1$ となる。

あてはまりのよさの指標である Adjusted R-squared をみても、それぞれ 0.6528 と 0.427 であり、平日と休日を区別せずに回帰をしたときの 0.2348 よりも改善されている。