

Crimtab Data : Tests of Normality

coop711

2018-03-31

Data Loading

```
load("./crimtab.RData")
str(crimtab_long_df)
```

```
## 'data.frame': 3000 obs. of 2 variables:
## $ finger: num 10 10.3 9.9 10.2 10.2 10.3 10.4 10.7 10 10.1 ...
## $ height: num 56 57 58 58 58 58 58 59 59 ...
```

Quetelet의 가슴둘레 자료에서 살핀 바와 같이 이 자료를 그대로 `ad.test` 등에 적용하면 매우 작은 p-value 가 예상됨.

```
# install.packages("nortest", repos = "https://cran.rstudio.com")
library(nortest)
# ad.test(crimtab_long_df[, 1])
# ad.test(crimtab_long_df[, 2])
kable(sapply(crimtab_long_df, ad.test))
```

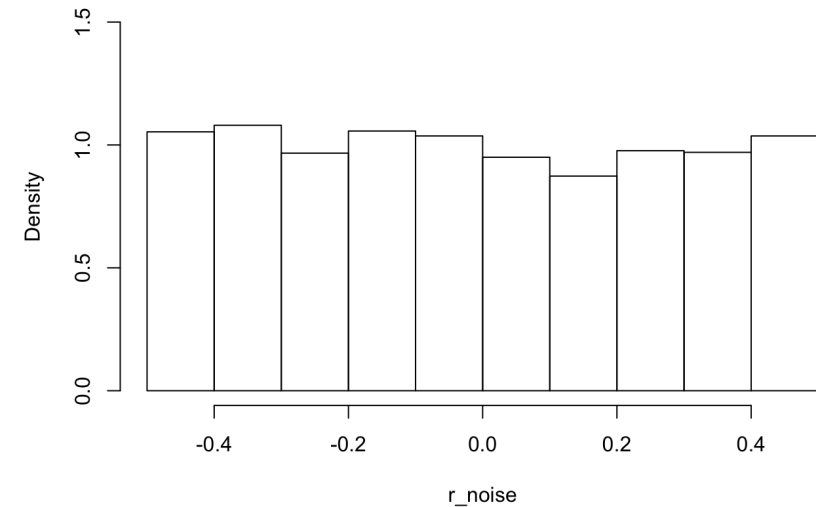
	finger	height
statistic	4.70938145946138	18.8367673791904
p.value	1.15288957655005e-11	3.7e-24
method	Anderson-Darling normality test	Anderson-Darling normality test
data.name	X[[1]]	X[[1]]

Random Noise

height 의 경우 인치 단위로 측정된 자료임. 이 점에 유의하여 원 자료의 모습에 가깝게 noise 를 넣기 위하여 생성

```
set.seed(1)
r_noise <- runif(3000) - 0.5
hist(r_noise,
     prob = TRUE,
     xlim = c(-0.5, 0.5),
     ylim = c(0, 1.5))
```

Histogram of r_noise



Plots

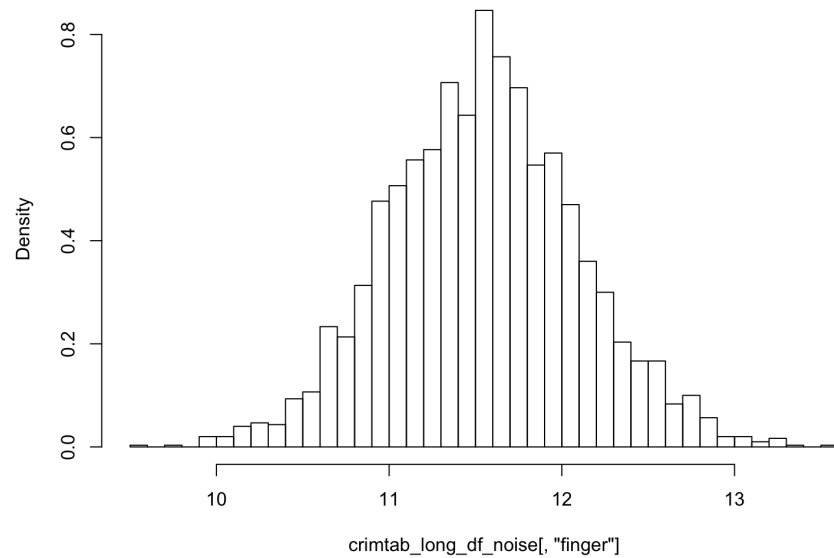
noise를 넣어 Histogram 과 QQnorm 작성

finger with noise

Histogram

```
crimtab_long_df_noise <- crimtab_long_df
crimtab_long_df_noise[, "finger"] <- crimtab_long_df[, "finger"] + r_noise/10
hist(crimtab_long_df_noise[, "finger"],
     prob = TRUE,
     nclass = 42)
```

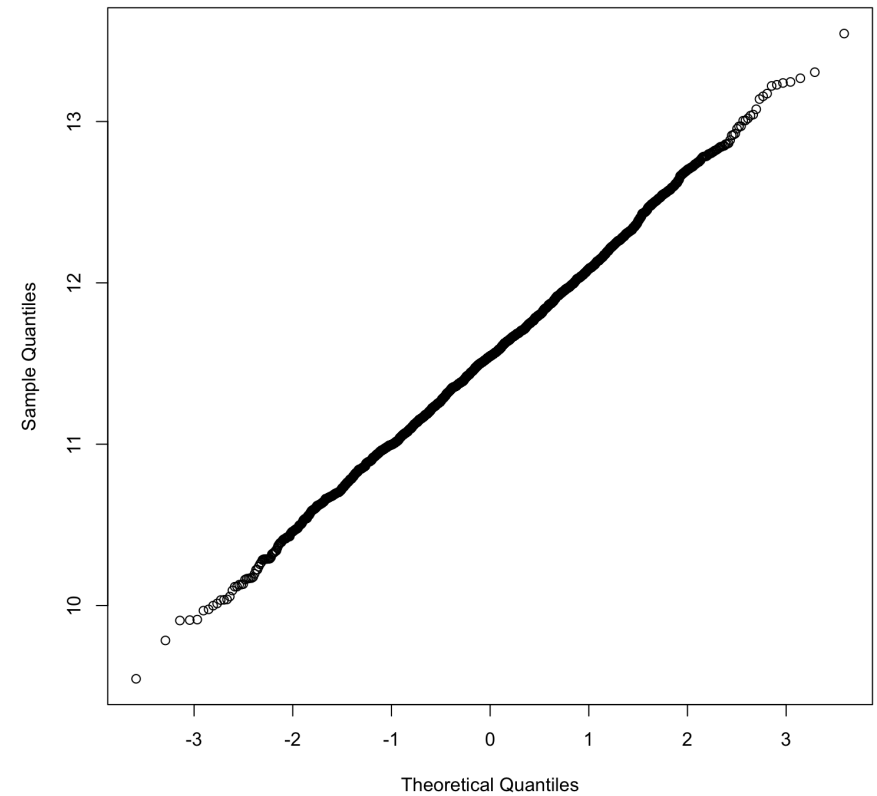
Histogram of crimtab_long_df_noise[, "finger"]



QQnorm

```
qqnorm(crimtab_long_df_noise[, "finger"])
```

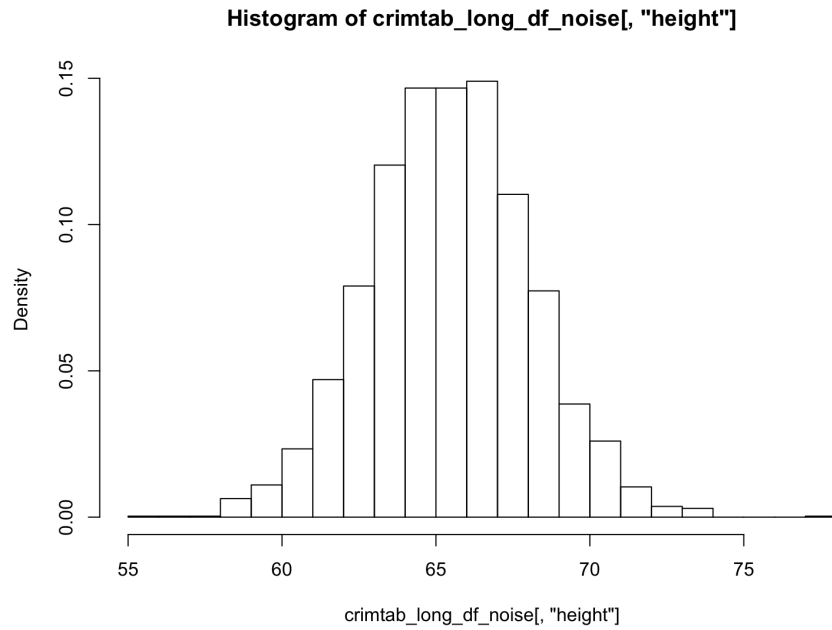
Normal Q-Q Plot



Height with noise

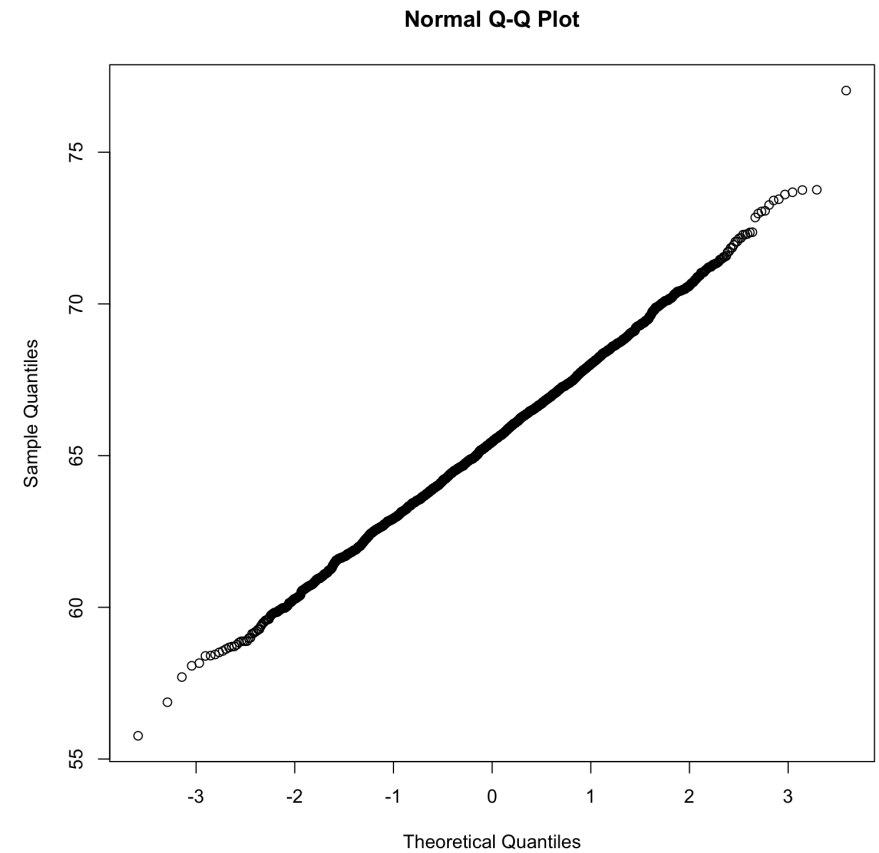
Histogram

```
crimtab_long_df_noise[, "height"] <- crimtab_long_df[, "height"] + r_noise
hist(crimtab_long_df_noise[, "height"],
     prob = TRUE,
     nclass = 22)
```



QQnorm

```
qqnorm(crimtab_long_df_noise[, "height"])
```



Tests of Normality

```
kable(sapply(crimtab_long_df_noise, ad.test))
```

	finger	height
statistic	0.689309974081425	0.330214935922413
p.value	0.0718113123018763	0.514030925598368
method	Anderson-Darling normality test	Anderson-Darling normality test
data.name	X[[i]]	X[[i]]

```
kable(sapply(crimtab_long_df_noise, cvm.test))
```

	finger	height
statistic	0.102810206983036	0.0443906432051351
p.value	0.102865272849146	0.600853915742615
method	Cramer-von Mises normality test	Cramer-von Mises normality test
data.name	X[[i]]	X[[i]]

```
kable(sapply(crimtab_long_df_noise, lillie.test))
```

	finger	height
statistic	0.0166259199968676	0.0107570113345079
p.value	0.0540195122352294	0.546641517923222
method	Lilliefors (Kolmogorov-Smirnov) normality test	Lilliefors (Kolmogorov-Smirnov) normality test
data.name	X[[i]]	X[[i]]

```
# ad.test(crimtab_long_df[, 2] + r_noise)
# cvm.test(crimtab_long_df[, 2] + r_noise)
# lillie.test(crimtab.long[,2]+r_noise)
# ad.test(crimtab.long[, 1] + r_noise/10)
# cvm.test(crimtab.long[, 1] + r_noise/10)
# lillie.test(crimtab.long[, 1] + r_noise/10)
```

```
save.image("./crimtab_test.RData")
```