

# Statistical Simulation

Chanseok Park (박찬석)

Applied Statistics Laboratory  
Department of Industrial Engineering  
Pusan National University

August 31, 2018

Hosted by SEC



부산대학교  
PUSAN NATIONAL UNIVERSITY

## 1 Definition of Statistical Simulation

## 2 Problems can be solved by statistical simulations

- Estimation of the probability when tossing a die or dice
- Ratio of the circumference of a circle
- Efficiency of estimators
- Breakdown point
- Correction factors ( $d_2$  and  $c_4$ )
- Hypothesis testing
- Development of `rt.test` R package
- Development of `weibullness` R package

## 1 Definition of Statistical Simulation

## 2 Problems can be solved by statistical simulations

- Estimation of the probability when tossing a die or dice
- Ratio of the circumference of a circle
- Efficiency of estimators
- Breakdown point
- Correction factors ( $d_2$  and  $c_4$ )
- Hypothesis testing
- Development of `rt.test` R package
- Development of `weibullness` R package

# Definition of Statistical Simulation

## Definition (wikipedia)

Monte Carlo methods (or Monte Carlo experiments) are a broad class of computational algorithms that rely on **repeated random sampling** to obtain numerical results. Their essential idea is using randomness to solve problems that might be **deterministic in principle**. They are often used in physical and mathematical problems and are most useful when it is difficult or impossible to use other approaches. *Monte Carlo methods are mainly used in three(actually much more?) problem classes*: optimization, numerical integration, and generating draws from a probability distribution.

- Monte Carlo methods  $\implies$  statistical or stochastic simulation.
- It can solve **stochastic** and **deterministic** problems.
- More applications in statistics: power, distribution, efficiency, bias-correction, etc.

# Tips for Simulation

From my teaching experience, many students are confused with:

- sample size  $n$  and
- the iteration (replication) number  $I$ .

- Sample size  $n$  is in general the number of random variable(s).
- The iteration number  $I$  is the number of repeated experiments.

# Problems can be solved by statistical simulations

In this talk, we will solve various problems by statistical simulations.

- 1 Estimation of the probability when tossing a die or dice.
- 2 The ratio of the circumference of a circle,  $\pi = 3.14\dots$
- 3 Efficiency of estimators.
- 4 Breakdown point.
- 5 Bias correction factors ( $d_2$  and  $c_4$ ) in Quality Control Chart.
- 6 Hypothesis testing.
- 7 R package (`rt.test`)
- 8 R package (`weibullness`)

# (1) Estimation of the probability

**We are interested in the probability when tossing a single die.**

## R Code for tossing a fair die

```
set.seed(1)
ITER = 10
x = sample(1:6, size=ITER, replace=TRUE)
table(x) / ITER
table( factor(x, levels=1:6)) / ITER # cosmetic
```

**What is the prob. of sum of two outcomes when tossing two dice.**

## R Code for tossing two dice (sum of two)

```
set.seed(1)
ITER = 10
x = sample(1:6, size=ITER, replace=TRUE)
y = sample(1:6, size=ITER, replace=TRUE)
table(x+y) / ITER
table( factor(x+y, levels=1:12)) / ITER # cosmetic
```

# (1) Estimation of the probability

Homework: 1. Find the probability of the difference of two dice.  
2. What if a die is not fair?



## (2) Ratio of the circumference of a circle, $\pi = 3.14 \dots$

The idea is very simple — consider a dart game.

The probability that a pin drops in a circle when a pin was thrown at random is given by:

$$\frac{\pi}{4} = \frac{\text{Area of unit circle}}{\text{Area of square}} \approx \frac{\# \text{ of pins in a circle}}{\# \text{ of all the pins}} = \hat{p}.$$

Thus, we have  $\pi \approx 4 \times \hat{p}$ .

◆ NB: Avoid using **for()** loop in R.

### R Code for finding $\pi = 3.14 \dots$

```
set.seed(1)
ITER = 100
x = runif(ITER, min=-1, max=1)
y = runif(ITER, min=-1, max=1)
4 * sum (x^2 + y^2 < 1) / ITER
```

Also, refer to Talk-1 at [▶ seminar](#).

(2) Ratio of the circumference of a circle,  $\pi = 3.14 \dots$

### (3) Efficiency of Estimators (mean, median, HL)

The relative efficiency of  $\hat{\theta}_2$  given  $\hat{\theta}_1$  is defined as

$$\text{RE}(\hat{\theta}_2, \hat{\theta}_1) = \frac{\text{Var}(\hat{\theta}_1)}{\text{Var}(\hat{\theta}_2)} \times 100\% \quad (1)$$

where  $\hat{\theta}_1$  is often a reference or baseline estimator (usually MLE).

Breakdown Points and AREs of the estimators under consideration.

	Location			Scale			
	Mean	Median	HL	SD	IQR	MAD	Shamos
ARE	100%	64%	96%	100%	38%	37%	86%
Breakdown	0%	50%	29%	0%	25%	50%	29%

- Refer to §2.2 of Lehmann Lehmann (1999), Rousseeuw and Croux (1993) and Serfling (2011); Staudte and Sheather (1990).
- $\text{HL} = \text{median}(X_i + X_j)/2$  for all  $i < j$  or  $i \leq j$ .

Refer to Talk-2 at [▶ seminar](#).

### (3) Efficiency of Estimators (mean, median, HL)

Here, we will focus on RE or ARE of median and HL with mean as a baseline.

#### R Code for Relative Efficiency

```
# Load Hodges-Lehmann and Shamos estimators.
source("https://raw.githubusercontent.com/AppliedStat/seminar/master/2018/R/Rsec.R")
ITER = 1000
n = 50
means = medians = HLs = numeric(ITER)

for ( i in 1:ITER ) {
  x = rnorm(n)
  means[i] = mean(x)
  medians[i] = median(x)
  HLs[i] = HL(x)
}

c( var(means) / var(medians), var(means) / var(HLs) )
```

◇NB: As  $n \rightarrow \infty$  and  $I \rightarrow \infty$ , RE converges to ARE.

## (4) Breakdown point

### Definition (Breakdown point)

The finite-sample breakdown point is the maximum proportion of incorrect observations (i.e. arbitrarily large observations) that an estimator can handle without leading to an egregiously incorrect estimate.

### R Code for Breakdown point

```
n = 100    # Number of observations
k = 20     # Number of contaminated obs.
x = c(runif(n-k), rep(Inf,k))
c(median(x), HL(x) )
k / n
```

- Idea: increase  $k$  until an estimator breaks down.
- Try  $k = 28, 29, 30, \dots, 49, 50, 51$ .

## (4) Breakdown point

### R Code for Breakdown point (a better code)

```
n = 100    # Number of observations
for ( k in 1: (n/2) ) {      # repeat with a different k
  x = c(runif(n-k), rep(Inf,k))
  print( c(k/n,median(x),HL(x)) )
}
```

- We obtained the finite-sample breakdown point.
- As increase  $n$  along with  $k$ , we can find the finite-sample breakdown point.
- ◇ The breakdown point of the median is  $1/2$  and that of the Hodges-Lehmann is  $1 - 1/\sqrt{2} = 0.29289$ .  
Also, refer to Talk-2 at ▶ seminar.

## (5-1) Bias correction factor ( $d_2$ )

The range,  $R = X_{(n)} - X_{(1)}$ , is not unbiased under the normality assumption. To make it unbiased, we adjust  $R$  with the correction factor  $d_2$  so that  $R/d_2$  is unbiased, that is,  $E[R/d_2] = \sigma$ . Then we have

$$d_2 = \frac{1}{\sigma} E[R].$$

Note that we have  $X = \sigma Z + \mu$ , where  $Z$  is from  $N(0, 1)$ . Thus, we have

$$d_2 = E[Z_{(n)} - Z_{(1)}] = 2 \int_0^{\infty} \left\{ 1 - [\Phi(z)]^n - [1 - \Phi(z)]^n \right\} dz,$$

where  $\Phi(\cdot)$  is the CDF of  $N(0, 1)$ .

Refer to Section 1 of [Talk-5-extra.pdf](#) at [seminar](#).

- To prove the above, it takes almost **ten pages**.
- $d_2$  depends on the sample size  $n$ .

## (5-1) Bias correction factor ( $d_2$ )

Idea for estimating  $E[Y]$ .

The theoretical expectation  $E[Y]$  is approximated by

$$\frac{1}{I} \sum_{i=1}^I Y_i.$$

Since  $d_2 = E[Z_{(n)} - Z_{(1)}] = E[Z_{(n)}] - E[Z_{(1)}]$ , we have

$$E[Z_{(n)}] \approx \frac{1}{I} \sum_{i=1}^I \max_i \quad \text{and} \quad E[Z_{(1)}] \approx \frac{1}{I} \sum_{i=1}^I \min_i,$$

where  $\max_i$  and  $\min_i$  are the maximum and minimum from a sample of size  $n$  at the  $i$ th iteration.



## (5-1) Bias correction factor ( $d_2$ )

### R Code for $d_2$

```
set.seed(1)
n = 10          # sample size
ITER = 1E4      # The larger, the more accurate d2
MAX = MIN = numeric(ITER)
for ( i in 1:ITER ) {
  Z = rnorm(n)
  MAX[i] = max(Z)
  MIN[i] = min(Z)
}
mean(MAX) - mean(MIN)
```

- Better R codes are provided in Section 1 of [Talk-5-extra.pdf](#) at [▶ seminar](#).
- For  $n = 10$ , we have  $d_2 = 3.078$  from the quality control textbook.
- The value through the simulation is given by  $d_2 = 3.092639$ .

## (5-2) Bias correction factor ( $c_4$ )

It is well known that the sample variance  $S^2$  is **unbiased** under the normal distribution assumption. That is

$$E[S^2] = \sigma^2,$$

where  $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)$ . However, the sample standard deviation  $S$  is **not** unbiased so that

$$E[S] \neq \sigma.$$

Most quality control textbooks use the correction factor,  $c_4$ , so that  $E[S/c_4] = \sigma$ . The correction factor  $c_4$  is theoretically derived as

$$c_4 = \sqrt{\frac{2}{n-1}} \cdot \frac{\Gamma(n/2)}{\Gamma(n/2 - 1/2)}.$$

Note that  $c_4$  is an unbiasing factor which is a function of sample size and it was originally used in ASTM E-11 (1976).

Refer to Section 2 of [Talk-5-extra.pdf](#) at [seminar](#).

## (5-2) Bias correction factor ( $c_4$ )

Since  $E[S/c_4] = \sigma$ , we have  $c_4 = \frac{1}{\sigma}E[S] = E[S/\sigma]$ . Note that  $S/\sigma$  is a **pivot**. Similar to the case of the range  $R$ , we have

$$c_4 = E[S_Z],$$

where  $S_Z$  is the sample standard deviation with a sample of size  $n$  from  $N(0, 1)$ .

### R Code for $d_2$

```
set.seed(1)
n = 10          # sample size
ITER = 1E4      # The larger, the more accurate d2
S = numeric(ITER)
for ( i in 1:ITER ) {
  Z = rnorm(n)
  S[i] = sd(Z)
}
mean(S)
```

## (5-3) Comparing $R$ , $R/d_2$ , $S$ , $S/c_4$

We have studied  $R$ ,  $R/d_2$ ,  $S$ , and  $S/c_4$ .

Then, are  $R/d_2$  and  $S/c_4$  better?

We usually check the bias and variance (or relative efficiency).

Let  $\hat{\theta}_i$  be the  $i$ th estimate from the simulation. Then the empirical bias and variance are given by

- Empirical bias:

$$\frac{1}{I} \sum_{i=1}^I \hat{\theta}_i - \theta.$$

- Empirical variance:

$$\frac{1}{I-1} \sum_{i=1}^I (\hat{\theta}_i - \bar{\hat{\theta}})^2.$$

◇NB: Based on the above variance, we can estimate the relative efficiency.

## (5-3) Comparing $R$ , $R/d_2$ , $S$ , $S/c_4$

### R Code for empirical bias and variance

```
# Load Hodges-Lehmann and Shamos estimators.
source("https://raw.githubusercontent.com/AppliedStat/seminar/master/2018/R/Rsec.R")
set.seed(1)
ITER = 1E5
R = Rd2 = SD = SDc4 = numeric(ITER)
for ( i in 1:ITER ) {
  X = rnorm(n)
  R[i] = diff(range(X))
  SD[i] = sd(X)
}
Rd2 = R / d2(n)
SDc4 = SD / c4(n)

# Bias of estimators
c( mean(R)-1, mean(Rd2)-1, mean(SD)-1, mean(SDc4)-1 )

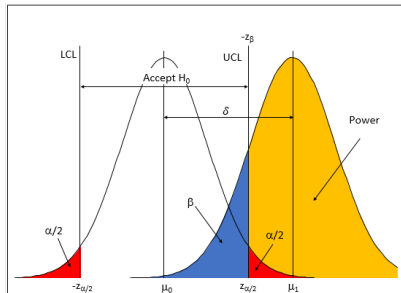
# Var. of estimators
c( var(R), var(Rd2), var(SD), var(SDc4) )
```

## (6) Hypothesis testing

Consider the statistical hypothesis testing:

$$H_0 : \mu = \mu_0 \text{ versus } H_1 : \mu \neq \mu_0.$$

		Decision	
		$H_0$	$H_1$
Truth	$H_0$	<b>OK</b> (coverage) ( $1 - \alpha$ )	<b>Type-I error</b> ( $\alpha$ )
	$H_1$	<b>Type-II error</b> ( $\beta$ )	<b>OK</b> (power) ( $K(\cdot) = 1 - \beta$ )



◇NB: Power is usually a function of  $\mu$  or  $\delta = \mu_1 - \mu_0$ .

## (6-1) Hypothesis testing (Example)

### Example 8.5-2 from Hogg et al. (2015)

Let  $X_1, X_2, \dots, X_{25}$  be a random sample from  $N(\mu, 100)$ . We want to test  $H_0 : \mu = 60$  versus  $H_1 : \mu > 60$ .

When we decided to **reject  $H_0$  if  $\bar{X} > 62$**  (critical/rejection region), find the power function of  $\mu$ .

Since  $n = 25$  and  $X_i \sim N(\mu, 100)$ , we have  $\bar{X} \sim N(\mu, 100/25)$ . Then  $Z = (\bar{X} - \mu)/\sqrt{100/25} = (\bar{X} - \mu)/2 \sim N(0, 1)$ . Thus, under  $H_1 : \mu$ , the power function is given by

$$\begin{aligned} K(\mu) &= P[\bar{X} > 62 \mid \mu] \\ &= P\left[\frac{\bar{X} - \mu}{2} > \frac{62 - \mu}{2}\right] = P\left[Z > \frac{62 - \mu}{2}\right] \\ &= 1 - \Phi\left[\frac{62 - \mu}{2}\right]. \end{aligned}$$

## (6-1) Hypothesis testing (Example: Power)

### Example 8.5-2 from Hogg et al. (2015)

The example state that the power at  $\mu = 65$  is  $K(65) = 0.9332 \dots$

### R Code for Example 8.5-2

```
# Theoretical power
K = function(mu) { 1-pnorm( (62-mu)/2 ) }

# Empirical power (through simulation)
set.seed(1);
ITER = 10000
n=25; sigma=10; mu=65
count = 0
for ( i in 1:ITER ) {
  X = rnorm(n=n, mean=mu, sd=sigma)
  # count if it is in a critical region.
  if (mean(X) > 62) count = count+1
}
c( count / ITER, K(65) )
```



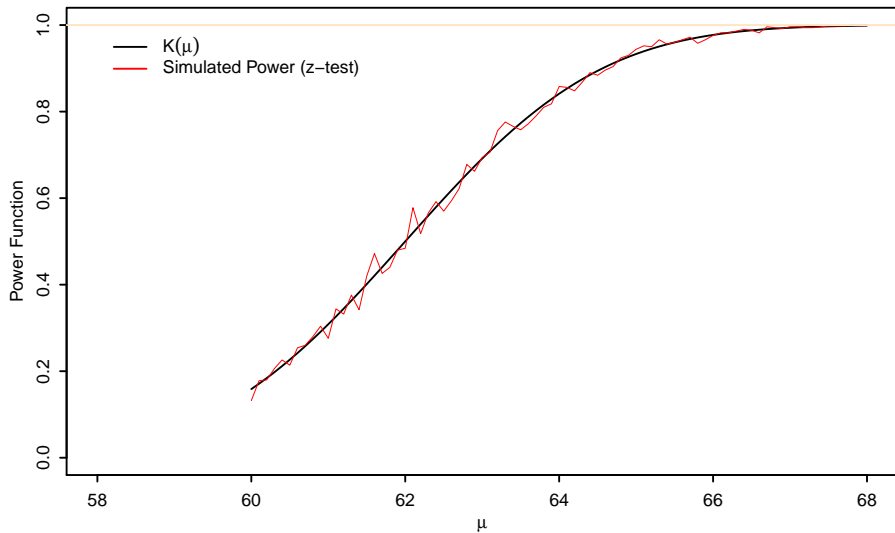
## (6-1) Hypothesis testing (Power function)

### R Code for Figure 8.5-2: Power function $K(\mu)$

```
# Empirical power (through simulation)
set.seed(1)
ITER = 1000
n=25; sigma=10; MU=seq(60, 68, length=81)
power = numeric(length(MU))
for ( j in 1:length(MU) ) {
  mu = MU[j]
  for ( i in 1:ITER ) {
    X = rnorm(n=n, mean=mu, sd=sigma)
    if ( mean(X) >= 62 ) power[j] = power[j] + 1/ITER
  }
}

# Compare the empirical power with the theoretical power
plot(MU, K(MU), xlim=c(58,68), ylim=c(0,1), type="l" )
lines(MU, power, col="red")
```

## (6-1) Hypothesis testing



## (6-2) Testing $H_0 : \mu = \mu_0$ vs. $H_0 : \mu \neq \mu_0$

- $\sigma$  is known (z-test): reject  $H_0$  if

$$|z| = \frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}} > z_{\alpha/2}$$

$$K_z(\mu) = 1 - \Phi\left(z_{\alpha/2} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right) + \Phi\left(-z_{\alpha/2} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right)$$

- $\sigma$  is **unknown** (t-test): reject  $H_0$  if

$$|t| = \frac{|\bar{x} - \mu_0|}{\textcolor{red}{S}/\sqrt{n}} > t_{\alpha/2}(n-1)$$

$$K_t(\mu) = 1 - \Phi_{\nu,\delta}(t_{\alpha/2}) + \Phi_{\nu,\delta}(-t_{\alpha/2}),$$

where  $\Phi_{\nu,\delta}(\cdot)$  is the CDF of the **non-central**  $t$ -distribution with  $\nu = n - 1$  degrees of freedom and non-centrality  $\delta = (\mu - \mu_0)/(\sigma/\sqrt{n})$ .

- ◆ Refer to Section 3 of Talk-5-extra.pdf at ▶ seminar  
 $\Rightarrow$  **4 pages!**

## (6-2) Hypothesis testing (Power function when $\sigma$ is known)

### R Code for empirical power of z-test ( $\sigma$ is known)

```
Kz = function(mu, alpha, mu0, sigma, n) {  
  z.cut = qnorm(1-alpha/2)  
  tmp = (mu-mu0)/(sigma/sqrt(n))  
  pnorm(z.cut + tmp, lower.tail=FALSE) + pnorm(-z.cut + tmp)  
}  
  
# Empirical Power  
set.seed(1); ITER = 1E4  
mu=1.5; mu0=0.5; alpha=0.05; sigma=1; n=5  
power = 0  
zcut = qnorm(1-alpha/2)  
for ( i in 1:ITER ) {  
  X = rnorm(n, mean=mu, sd=1)  
  if(abs(mean(X)-mu0)/(sigma/sqrt(n)) > zcut) power=power+1/ITER  
}  
  
# Compare  
c( Kz(mu, alpha, mu0, sigma, n), power )
```

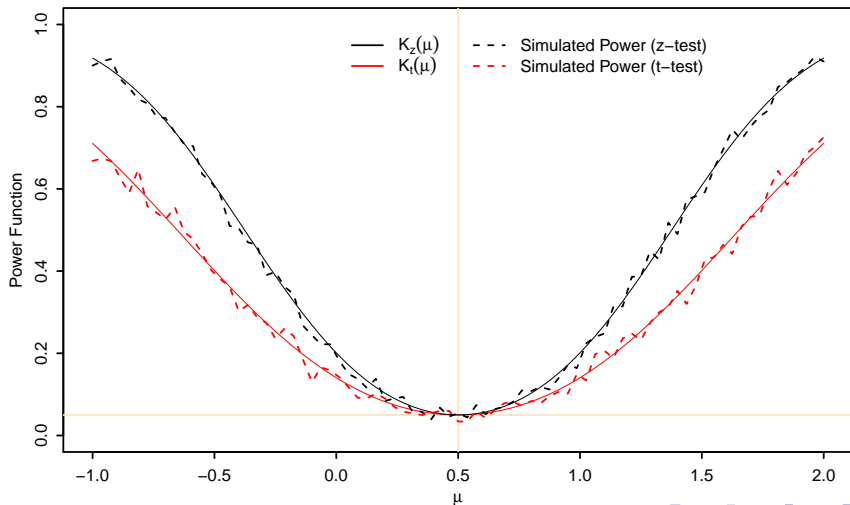
## (6-3) Hypothesis testing (Power when $\sigma$ is unknown)

### R Code for empirical power of $t$ -test ( $\sigma$ is unknown)

```
Kt = function(mu, alpha, mu0, sigma, n) {  
  t.cut = qt(1-alpha/2,df=n-1)  
  ncp = (mu-mu0)/(sigma/sqrt(n))  
  pt(t.cut,df=n-1,ncp=ncp,lower.tail=F)+pt(-t.cut,df=n-1,ncp=ncp)  
}  
  
# Empirical Power  
set.seed(1); ITER = 1E4  
mu=1.5; mu0=0.5; alpha=0.05; sigma=1; n=5  
power = 0  
tcut = qt(1-alpha/2, df=n-1) # tcut is used instead of zcut  
for ( i in 1:ITER ) {  
  X = rnorm(n, mean=mu, sd=1)  
  # sd(X) is used instead of known sigma  
  if(abs(mean(X)-mu0)/(sd(X)/sqrt(n)) > tcut) power=power+1/ITER  
}  
  
# Compare  
c( Kt(mu, alpha, mu0, sigma, n), power )
```

## (6-3) Hypothesis testing: Power

Testing  $H_0 : \mu = 1/2$  versus  $H_1 : \mu \neq 1/2$  with  $\alpha = 0.05$ ,  $n = 5$ , and  $\sigma = 1$ . Refer to Section 3.3 of [Talk-5-extra.pdf](#) at [seminar](#).



## (7-1) Warming-up before rt.test R package

### $t$ -test statistic

$$T = \frac{\bar{X} - \mu}{S_X / \sqrt{n}}, \quad (2)$$

where  $\bar{X} = \sum X_i / n$ ,  $S_X^2 = \sum (X_i - \bar{X})^2 / (n - 1)$  and  $X_i$  are from  $N(\mu, \sigma^2)$ . Substituting  $X_i = \sigma Z_i + \mu$  into (2), we have

$$T = \frac{\bar{Z} - \mu}{S_Z / \sqrt{n}},$$

Note that  $T$  is a **pivot**.

- Assume that we know **only** the normal distribution.
- How to know **the distribution** of  $T$  above.

## (7-1) Warming-up before `rt.test` R package

R Code for the critical value of the two-side  $t$ -test when  $\alpha = 5\%$

```
set.seed(1)
ITER = 1.0E4; n=6;
Tstat = numeric(ITER)
for ( i in 1:ITER ) {
  Z = rnorm(n)
  Tstat[i] = mean(Z) / (sd(Z)/sqrt(n))
}

# Critical values
quantile(Tstat, prob=0.975)
qt(0.975, df=n-1)
qnorm(0.975) # 1.96

# Plot
hist(Tstat,breaks=100,probability=T,xlim=c(-5,5),ylim=c(0,0.4))
lines( density(Tstat))
curve( dnorm(x), xlim=c(-5,5), add=T, col="red" )
```



## (7-2) rt.test R package

### Recall a basic idea

$$T = \frac{\boxed{\bar{X}} - \mu}{\boxed{S} / \sqrt{n}} \quad (3)$$

- $\boxed{\bar{X}}$  is replaced by median or Hodges and Lehmann (1963).
- $\boxed{S}$  is replaced by MAD or Shamos (1976).

- **Is it enough?** No.
- **Then, is it distributed as  $N(0, 1)$  or  $t$ -distribution?** Neither.  
 $T_A$  or  $T_B \overset{\circ}{\sim} N(0, 1)$  (Refer to KIIE at Seoul, November 2018).  
Thus,  $T_A$  or  $T_B$  works for a large sample.  
**What if a sample is small?**

◇ NB: Refer to Talk-2 at [▶ seminar](#) and `rt.test` R package at [▶ rt.test](#).

## (7-2) rt.test R package

### Recall $T_A$ and $T_B$

Refer to Park and Wang (2018a) at [ArXiv](#).

$$T_A = \sqrt{\frac{2n}{\pi}} \Phi^{-1}\left(\frac{3}{4}\right) \cdot \frac{\operatorname{median}_{1 \leq i \leq n} X_i - \mu}{\operatorname{median}_{1 \leq i \leq n} |X_i - \operatorname{median}_{1 \leq i \leq n} X_i|} \xrightarrow{d} N(0, 1), \quad (4)$$

where  $\Phi^{-1}(\cdot)$  is the inverse of the standard normal cumulative distribution function and  $\xrightarrow{d}$  denotes convergence in distribution.

$$T_B = \sqrt{\frac{6n}{\pi}} \Phi^{-1}\left(\frac{3}{4}\right) \frac{\operatorname{median}_{i \leq j} \left( \frac{X_i + X_j}{2} \right) - \mu}{\operatorname{median}_{i \leq j} (|X_i - X_j|)} \xrightarrow{d} N(0, 1). \quad (5)$$

◇NB:  $T_A$  and  $T_B$  are **pivotal**. See Park (2018a) and Jeong et al. (2018).

## (7-2) rt.test R package

### R Code for the critical value of rt.test R package

```
set.seed(1)
ITER = 1.0E4; n=6;
TA = numeric(ITER)
for ( i in 1:ITER ) {
  Z = rnorm(n)
  TA[i] = sqrt(2*n/pi)*qnorm(3/4)*median(Z)/mad(Z,constant=1)
}

# Install rt.test package
install.packages("rt.test")
library("rt.test")

# Critical values
quantile(TA, prob=0.975)
q.robustified.t(0.975, n=n) # This Package used ITER = 1.0E8

qnorm(0.975) # 1.96 (compare it with the above)
```

## (8) weibullness R package

Recall: Weibull distribution (reparametrized form)

$$F(x) = 1 - \exp(-\lambda x^\alpha) \text{ and } f(x) = \lambda \alpha x^{\alpha-1} \exp(-\lambda x^\alpha),$$

It is immediate from the Weibull CDF  $F(x) = p$  that

$$\log(1 - p) = -\lambda x_p^\alpha.$$

It follows that

$$\log \{ -\log(1 - p) \} = \log \lambda + \alpha \log x_p.$$

This implies that the plot of

$$\log \{ -\log(1 - p) \} \text{ versus } \log x_p$$

draws a straight line with the slope  $\alpha$  and the intercept  $\log \lambda$ .

◇NB: Refer to Park (2017) and [Talk-3](#) at [seminar](#) and weibullness R package at [weibullness](#).

## (8) weibullness R package

Recall: How to determine whether the data are from Weibull or not

$H_0$  : Weibull versus  $H_1$  : non-Weibull.

- Recall:  $\log \{ -\log(1 - p) \} = \log \lambda + \alpha \log x_p$ .
- Thus  $\log \{ -\log(1 - \hat{p}_i) \}$  versus  $\log x_{(i)}$  draws a straight line, where  $\hat{p}_i = (i - 0.375)/(n + 0.25)$ . Thus, the linearity measure (sample correlation) can be used for Weibullness test.
- Idea: Large  $r$  implies Weibullness.

$$H_0 : r \geq r_0 \text{ versus } H_1 : r < r_0$$

Then, how small is small enough for  $r$  to **reject**  $H_0$ ?

- To this end, we should know the distribution of  $r$  and then find the critical value.
- ◇ NB:  $r$  is a **pivot**. Thus, for Weibull with any parameter values, the results are the same.  $\implies$  Use **exponential distribution with rate one**.

## (8) weibullness R package

### Algorithm

- Step-1: Generate Weibull random observations of size  $n$  from any Weibull. Note Weibull(1,1) is an exponential distribution.
- Step-2: Sort the data. Denote  $x_{(i)}$ .
- Step-3: Calculate  $\log \{ -\log(1 - \hat{p}_i) \}$ ,  
where  $\hat{p}_i = (i - 0.375)/(n + 0.25)$  or  $\hat{p}_i = (i - 0.5)/n$ .
- Step-4: Calculate the sample correlation between  
 $\log \{ -\log(1 - \hat{p}_i) \}$  and  $\log x_{(i)}$ .
- Step-5: Repeat the above (say, up to  $I$  iteration numbers).
- Step-6: Find the empirical quantiles for critical values (say, 5%, 10%, etc.)
- ◇ NB: Compare the sample correlation with the empirical quantile (critical value) in order to make a decision.

## (8) weibullness R package

### R Code for the critical value of weibullness R package

```
set.seed(1)
ITER = 1.0E4; n=23;
cors = numeric(ITER)
for ( i in 1:ITER ) {
  x = sort( rexp(n) )      # Step 1-2
  p = ppoints(n)          # Step 3
  cors[i] = cor(log(-log(1-p)),log(x))  # Step 4
} # Repeat (i=1,2,...,ITER) # Step 5

alphas = c(0.01,0.02,0.025,0.5,0.1,0.2)
quantile(cors, prob=alphas) # Step 6

# Install weibullness package
install.packages("weibullness")
library("weibullness")
wp.test.critical(alphas,n) # This package used ITER=1E8
```

# References

- ASTM E-11 (1976). Manual on Presentation of Data and Control Chart Analysis. American Society for Testing and Materials, Philadelphia, PA, 4th edition.
- Hodges, J. L. and Lehmann, E. L. (1963). Estimates of location based on rank tests. Annals of Mathematical Statistics, 34:598–611.
- Hogg, R. V., Tanis, E. A., and Zimmerman, D. L. (2015). Probability and Statistical Inference. Pearson, 9th edition.
- Jeong, R., Son, S. B., Lee, H. J., and Kim, H. (2018). On the robustification of the z-test statistic. Presented at KIIE Conference, Gyeongju, Korea. April 6, 2018.
- Lehmann, E. L. (1999). Elements of Large-Sample Theory. Springer, New York.



# References

- Park, C. (2017). Weibullness test and parameter estimation of the three-parameter Weibull model using the sample correlation coefficient. International Journal of Industrial Engineering: Theory, Applications and Practice, 24:376–391.
- Park, C. (2018a). Note on the robustification of the Student  $t$ -test statistic using the median and the median absolute deviation. <https://arxiv.org/abs/1805.12256>. ArXiv e-prints.
- Park, C. (2018b). weibullness: Goodness-of-fit test for Weibull (Weibullness test). <https://CRAN.R-project.org/package=weibullness>. R package version 1.18.6.
- Park, C. and Wang, M. (2018a). Empirical distributions of the robustified  $t$ -test statistics. <https://arxiv.org/abs/1807.02215>. ArXiv e-prints.

# References

- Park, C. and Wang, M. (2018b). `rt.test`: Robustified t-test.  
<https://CRAN.R-project.org/package=rt.test>. R package version 1.18.7.9.
- Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation. Journal of the American Statistical Association, 88:1273–1283.
- Serfling, R. J. (2011). Asymptotic relative efficiency in estimation. In Lovric, M., editor, Encyclopedia of Statistical Science, Part I, pages 68–82. Springer-Verlag, Berlin.
- Shamos, M. I. (1976). Geometry and statistics: Problems at the interface. In Traub, J. F., editor, Algorithms and Complexity: New Directions and Recent Results, pages 251–280. Academic Press, New York.
- Staudte, R. G. and Sheather, S. J. (1990). Robust Estimation and Testing. John Wiley & Sons, New York.