BAB 5

DISTRIBUSI SAMPLING

Distribusi Populasi (Population Distribution)

Distribusi populasi dapat diartikan sebagai distribusi probabilitas dari data populasi. Andaikan dalam suatu kelas hanya terdiri lima mahasiswa jurusan matematika. Berikut disajikan nilai ujian matakuliah kalkulus dari lima mahasiswa tersebut.

70, 75, 80, 80, 90

Andaikan X menyatakan nilai ujian matakuliah kalkulus dan $P(X = x)$ atau $f(x)$ menyatakan probabilitas dari suatu nilai ujian matakuliah kalkulus. Berikut disajikan distribusi probabilitas dari data populasi nilai ujian matakuliah kalkulus (Tabel 5.1).

Tabel 5.1 Distribusi Probabilitas dari Data Populasi Nilai Ujian Kalkulus

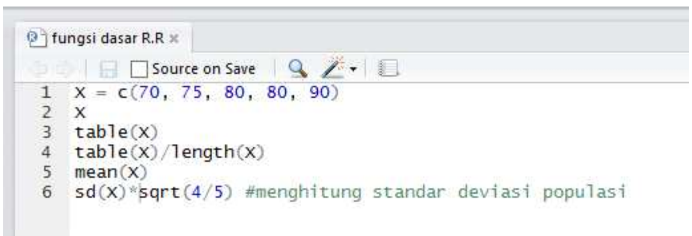
X	$P(X = x)$
70	0.2
75	0.2
80	0.4
90	0.2
$\sum P(X = x) = \sum f(x) = 1$	

Nilai rata-rata dan standar deviasi berdasarkan data pada Tabel 5.1 dihitung sebagai berikut.

$$\mu = \frac{70 + 75 + 80 + 80 + 90}{5} = 79$$

$$\sigma = \sqrt{\frac{(70 - 79)^2 + (75 - 79)^2 + \cdots + (90 - 79)^2}{5}} = 6{,}633$$

Perhatikan bahwa μ dan σ merupakan nilai-nilai parameter populasi. Parameter dapat diartikan sebagai suatu nilai atau ukuran yang dihitung berdasarkan populasi. Gambar 5.1 dan Gambar 5.2 merupakan ilustrasi dalam R.



Gambar 5.1

```
X = c(70, 75, 80, 80, 90)
X
```

```
## [1] 70 75 80 80 90
```

```
table(X)
```

```
## X
## 70 75 80 90
##  1  1  2  1
```

Untuk menampilkan distribusi probabilitas.

```
table(X)/length(X)
```

```
## X
##   70   75   80   90
## 0.2 0.2 0.4 0.2
```

```
mean(X)
```

```
## [1] 79
```

Penambahan **sqrt(4/5)** dengan maksud untuk menghitung standar deviasi populasi. Jika **sqrt(4/5)** dihilangkan, berarti menghitung standar deviasi sampel (bukan populasi).

```
sd(X)*sqrt(4/5) #menghitung standar deviasi populasi
```

```
## [1] 6.63325
```

Gambar 5.2

Distribusi Sampling Rata-Rata Sampel $\bar{X}$ (Sampling Distribution of $\bar{X}$)

Berbeda dengan statistika deskriptif yang rangkaian pengerjaannya meliputi mengorganisasi (*organizing*), menampilkan (*displaying*), dan menjelaskan data dengan menggunakan tabel, grafik, serta ukuran-ukuran seperti rata-rata, median, serta modus, pada statistika inferensi sampai pada tahap pengambilan keputusan atau prediksi mengenai populasi berdasarkan sampel yang diteliti. Konsep mengenai **distribusi sampling** memberikan teori yang penting untuk membuat prosedur-prosedur statistik inferensi. Daniel (2005:129) menyatakan sebagai berikut.

"*Sampling distributions serve two purposes: (1) they allow us to answer probability questions about sample statistics, and (2) they provide the necessary theory for making statistical inference procedures valid*".

Nilai dari parameter suatu populasi bersifat konstan. Dalam hal ini, untuk setiap data populasi hanya memiliki satu nilai rata-rata populasi μ. Namun hal ini belum tentu berlaku untuk rata-rata sampel $\bar{X}$. Sampel-sampel yang ditarik dari populasi yang sama dan dengan ukuran yang sama dapat menghasilkan nilai rata-rata sampel yang berbeda-beda. Jadi, nilai rata-rata sampel bergantung pada nilai-nilai yang berada dalam sampel tersebut. **Oleh karena itu, rata-rata sampel $\bar{X}$ merupakan variabel acak (*random variable*)**. Sebagaimana pada variabel acak, **maka rata-rata sampel $\bar{X}$ memiliki distribusi probabilitas**. Distribusi probabilitas $\bar{X}$ sering disebut dengan istilah **distribusi sampling dari $\bar{X}$**. Ukuran-ukuran statistik lainnya seperti median, modus, dan standar deviasi juga memiliki distribusi sampling (Mann dan Lacke, 2011:302).

Pada pembahasan sebelumnya mengenai "Distribusi Probabilitas", diketahui data populasi sebagai berikut.

70, 75, 80, 80, 90

Andaikan masing-masing nilai diberi kode huruf sebagai berikut.

V = 70, W = 75, X = 80, Y = 80, dan Z = 90

Maka, V, W, X, Y, dan Z merupakan kode-kode huruf yang menyatakan kelima nilai ujian matakuliah kalkulus. Kemudian misalkan akan diambil sampel yang terdiri tiga nilai tanpa pengembalian (*without replacement*). Maka banyaknya kemungkinan sampel yang terambil sebagai berikut.

$$C_3^5 = \frac{5!}{(5-3)!\,3!} = \frac{5.4.3.2.1}{(2.1)(3.2.1)} = 10\ kemungkinan\ sampel$$

VWX, VWY, VWZ, VXY, VXZ, VYZ, WXY, WXZ, WYZ, XYZ

Tabel 5.3 Sampel-Sampel yang Mungkin Terambil beserta Nilai Rata-Rata

Sampel	Nilai-Nilai dalam Sampel			$\bar{X}$
VWX	70	75	80	75
VWY	70	75	80	75
VWZ	70	75	90	78.33
VXY	70	80	80	76.67
VXZ	70	80	90	80
VYZ	70	80	90	80
WXY	75	80	80	78.33
WXZ	75	80	90	81.67
WYZ	75	80	90	81.67
XYZ	80	80	90	83.33

Perhatikan bahwa terdapat 10 kemungkinan sampel. Sampel VWX berarti mengandung nilai 70, 75, dan 80, sampel WYZ berarti mengandung nilai 75, 80, dan 90, dan seterusnya. Tabel 5.3 menyajikan sampel-sampel yang mungkin terambil beserta penghitungan nilai rata-rata. Berdasarkan Tabel 5.3, selanjutnya dibentuk tabel distribusi frekuensi dan frekuensi relatif berdasarkan nilai rata-rata sampel (Tabel 5.4). Tabel 5.5 menyajikan distribusi sampling dari rata-rata sampel $\bar{X}$ berdasarkan data pada Tabel 5.3.

Tabel 5.5 menyajikan distribusi probabilitas dari rata-rata sampel $\bar{X}$. Sebagai contoh probabilitas untuk memperoleh sampel yang memiliki nilai rata-rata 76,67 sebesar 0,2. Atau dapat dinyatakan sebagai berikut.

$$P(\bar{X} = 81.67) = 0.20$$

Tabel 5.4 Distribusi Frekuensi dan Frekuensi Relatif Berdasarkan Nilai Rata-Rata Sampel

$\bar{X}$	Frekuensi	Frekuensi Relatif
75	2	0.2
76.67	1	0.1
78.33	2	0.2
80	2	0.2
81.67	2	0.2
83.33	1	0.1
Jumlah	10	1

Tabel 5.5 Distribusi Sampling dari $\bar{X}$ dengan Ukuran Sampel sebanyak 3

$\bar{X}$	$P(\bar{X}=\bar{x})=f(\bar{x})$
75	0.2
76.67	0.1
78.33	0.2
80	0.2
81.67	0.2
83.33	0.1
$\sum P(\bar{X}=\bar{x})=1$	

Berikut diberikan ilustrasi dalam R.

```
fungsi dasar R.R
1  x = c(70, 75, 80, 80, 90)
2  x
3  library(prob)
4  urnsamples(c(70,75,80,80,90), size = 3, replace = FALSE, ordered = FALSE)
5
6
7  x = c(70, 75, 80, 80, 90)
8  x
9  sampel = combn(x, 3) #pengambilan sampel tanpa pengembalian dan tanpa memperhatikan urutan
10 sampel
11 ratarata = colMeans(sampel)
12 ratarata=format(ratarata, digits=4) #pengaturan desimal
13 prop.table(table(ratarata))
14 table(ratarata)/length(ratarata)
15 barplot(table(ratarata))
```

Gambar 5.3

Pada Gambar 5.3 mengaktifkan *package* **prob** (kode R baris 3) dengan maksud untuk menggunakan fungsi **urnsamples**.

```
x = c(70, 75, 80, 80, 90)
x
```

```
## [1] 70 75 80 80 90
```

```
library(prob)
```

Gambar 5.4

```
urnsamples(c(70,75,80,80,90), size = 3, replace = FALSE, ordered = FALSE)
```

```
##    X1 X2 X3
## 1  70 75 80
## 2  70 75 80
## 3  70 75 90
## 4  70 80 80
## 5  70 80 90
## 6  70 80 90
## 7  75 80 80
## 8  75 80 90
## 9  75 80 90
## 10 80 80 90
```

Gambar 5.5

Pada Gambar 5.5, penggalan kode R **replace = FALSE** berarti pengambilan sampel tanpa pengembalian, serta pada penggalan kode R **ordered = FALSE** berarti tanpa memperhatikan urutan.

```
sampel = combn(X, 3) #pengambilan sampel tanpa pengembalian dan tanpa memperhatikan urutan
sampel
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
## [1,]   70   70   70   70   70   70   75   75   75    80
## [2,]   75   75   75   80   80   80   80   80   80    80
## [3,]   80   80   90   80   90   90   80   90   90    90
```

Gambar 5.6

Pada Gambar 5.6 menyajikan alternatif kode R (dari yang sebelumnya) untuk menampilkan seluruh kemungkinan sampel yang mungkin terambil. Pada Gambar 5.6 menggunakan fungsi **combn** (*combination*).

```
ratarata = colMeans(sampel)
ratarata=format(ratarata, digits=4) #pengaturan desimal
prop.table(table(ratarata))
```

```
## ratarata
## 75.00 76.67 78.33 80.00 81.67 83.33
##   0.2   0.1   0.2   0.2   0.2   0.1
```

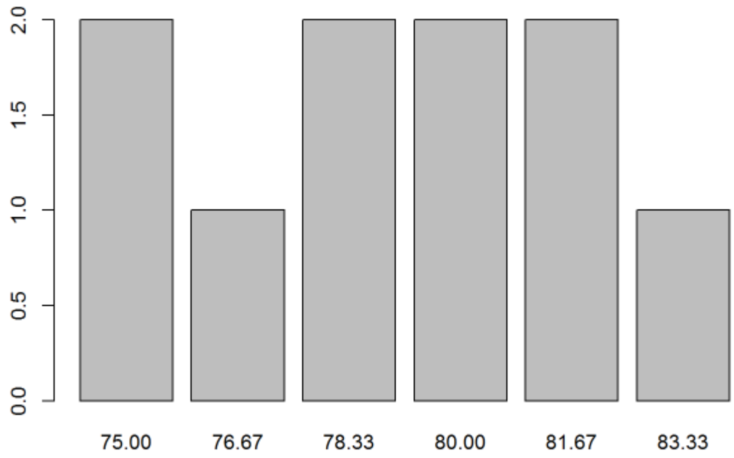
```
table(ratarata)/length(ratarata)
```

```
## ratarata
## 75.00 76.67 78.33 80.00 81.67 83.33
##   0.2   0.1   0.2   0.2   0.2   0.1
```

Gambar 5.7

Pada Gambar 5.7 menyajikan distribusi probabilitas dari rata-rata sampel $\bar{X}$. Pada Gambar 5.8 menyajikan grafik batang yang menyajikan frekuensi dari setiap nilai rata-rata sampel $\bar{X}$. Berdasarkan Gambar 5.8, nilai rata-rata 75 sebanyak 2, nilai rata-rata 76,67 sebanyak 1, dan seterusnya.

```
barplot(table(ratarata))
```



Gambar 5.8

Rata-Rata dari Distribusi Sampling Rata-Rata Sampel $\bar{X}$

Rata-rata dari distribusi sampling $\bar{X}$ (*mean of the sampling distribution of* $\bar{X}$) atau rata-rata dari $\bar{X}$ dilambangkan dengan $\mu_{\bar{X}}$. Berdasarkan Tabel 5.3, berikut akan dihitung rata-rata dari distribusi sampling $\bar{X}$ serta rata-rata populasinya.

$$\mu_{\bar{X}} = \frac{75 + 75 + 78{,}33 + \cdots + 83{,}33}{10} = 79$$

$$\mu = \frac{70 + 75 + 80 + 80 + 90}{5} = 79$$

Perhatikan bahwa berdasarkan perhitungan diperoleh $\mu_{\bar{X}} = 79$ dan $\mu = 79$**.** Mann dan Lacke (2011:307) menyatakan sebagai berikut.

"*The mean of the sampling distribution of* $\bar{X}$ ***is always equal*** *to the mean of the population. Thus,* $\mu_{\bar{X}} = \mu$".

Rata-rata sampel $\bar{X}$ disebut juga sebagai *estimator* atau penduga terhadap rata-rata populasi μ. Suatu statistik dikatakan sebagai estimator tak-bias atau *unbiased estimator* jika nilai rata-rata dari distribusi sampling statistik tersebut sama dengan nilai parameter tertentu. Perhatikan bahwa statistik rata-rata sampel $\bar{X}$ merupakan estimator tak-bias dari parameter rata-rata populasi (μ), karena nilai rata-rata dari distribusi sampling rata-rata $\bar{X}$ selalu sama dengan rata-rata populasi, yakni

$$\mu_{\bar{X}} = \mu.$$

Berikut diberikan ilustrasi dalam R.

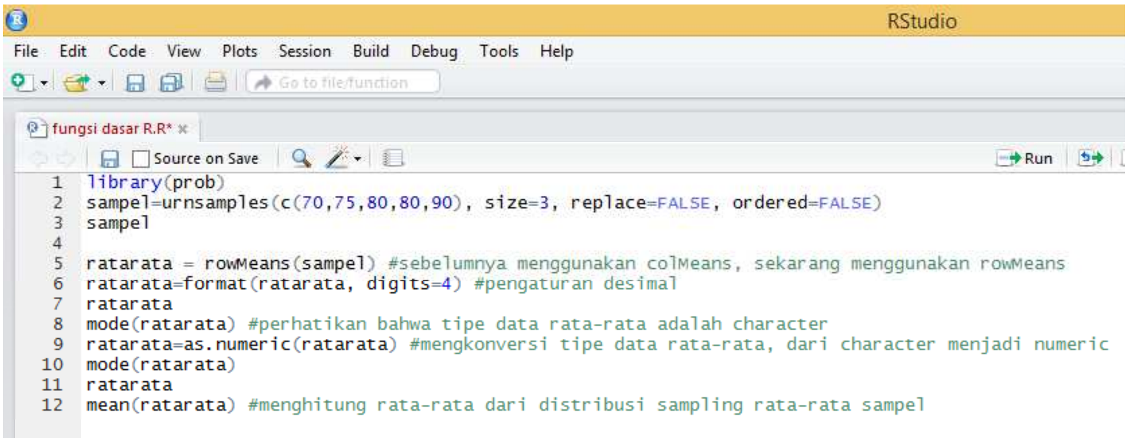
```
1  X = c(70, 75, 80, 80, 90)
2  X
3  sampel = combn(X, 3) #pengambilan sampel tanpa pengembalian dan tanpa memperhatikan urutan
4  ratarata = colMeans(sampel)
5  ratarata=format(ratarata, digits=4) #pengaturan desimal
6  ratarata
7  mode(ratarata) #perhatikan bahwa tipe data rata-rata adalah character
8  ratarata=as.numeric(ratarata) #mengkonversi tipe data rata-rata, dari character menjadi numeric
9  mode(ratarata)
10 ratarata
11 mean(ratarata) #menghitung rata-rata dari distribusi sampling rata-rata sampel
```

Gambar 5.9

```
X = c(70, 75, 80, 80, 90)
X

## [1] 70 75 80 80 90

sampel = combn(X, 3) #pengambilan sampel tanpa pengembalian dan tanpa memperhatikan urutan
ratarata = colMeans(sampel)
ratarata=format(ratarata, digits=4) #pengaturan desimal
ratarata

##  [1] "75.00" "75.00" "78.33" "76.67" "80.00" "80.00" "78.33" "81.67"
##  [9] "81.67" "83.33"

mode(ratarata) #perhatikan bahwa tipe data rata-rata adalah character

## [1] "character"

ratarata=as.numeric(ratarata) #mengkonversi tipe data rata-rata, dari character menjadi numeric
mode(ratarata)

## [1] "numeric"

ratarata

##  [1] 75.00 75.00 78.33 76.67 80.00 80.00 78.33 81.67 81.67 83.33

mean(ratarata)

## [1] 79
```

Gambar 5.10

Berdasarkan Gambar 5.9, kode R pada baris 7 bertujuan untuk mengetahui tipe atau jenis data dari variabel **ratarata**. Sementara kode R pada baris 8 bertujuan untuk mengkonversi jenis data variabel **ratarata**, dari *character* menjadi *numeric*. Setelah dikonversi menjadi *numeric*, barulah bisa dihitung nilai rata-rata dari distribusi sampling rata-rata sampel (kode R pada baris 11). Berikut alternatif kode R untuk memperoleh rata-rata dari distribusi sampling rata-rata sampel (perhatikan Gambar 5.11 sampai dengan Gambar 5.13).



Gambar 5.11

```
sampel=urnsamples(c(70,75,80,80,90), size=3, replace=FALSE, ordered=FALSE)
sampel
```

```
##    X1 X2 X3
## 1  70 75 80
## 2  70 75 80
## 3  70 75 90
## 4  70 80 80
## 5  70 80 90
## 6  70 80 90
## 7  75 80 80
## 8  75 80 90
## 9  75 80 90
## 10 80 80 90
```

```
ratarata = rowMeans(sampel)
ratarata=format(ratarata, digits=4) #pengaturan desimal
ratarata
```

```
##  [1] "75.00" "75.00" "78.33" "76.67" "80.00" "80.00" "78.33" "81.67"
##  [9] "81.67" "83.33"
```

```
mode(ratarata) #perhatikan bahwa tipe data rata-rata adalah character
```

```
## [1] "character"
```

```
ratarata=as.numeric(ratarata) #mengkonversi tipe data rata-rata, dari character menjadi numeric
mode(ratarata)
```

```
## [1] "numeric"
```

```
ratarata
```

```
##  [1] 75.00 75.00 78.33 76.67 80.00 80.00 78.33 81.67 81.67 83.33
```

```
mean(ratarata) #menghitung rata-rata dari distribusi sampling rata-rata sampel
```

```
## [1] 79
```

Gambar 5.12

Standar Deviasi dari Distribusi Sampling Rata-Rata Sampel $\bar{X}$

Diketahui pada pembahasan sebelumnya bahwa rata-rata dari distribusi sampling rata-rata $\bar{X}$ dilambangkan dengan simbol $\mu_{\bar{X}}$, sedangkan rata-rata populasi dilambangkan dengan simbol μ. Standar deviasi dari distribusi sampling rata-rata $\bar{X}$ dilambangkan dengan simbol $\sigma_{\bar{X}}$, sedangkan standar deviasi populasi dilambangkan dengan simbol σ. Pada pembahasan sebelumnya diketahui bahwa rata-rata dari distribusi sampling rata-rata $\bar{X}$ sama dengan rata-rata populasi μ, yakni

$$\mu_{\bar{X}} = \mu.$$

Namun pada standar deviasi dari distribusi sampling rata-rata $\bar{X}$ tidak sama dengan standar deviasi populasi (kecuali jika $n = 1$). Sebagai contoh untuk kasus $n = 1$, misalkan suatu populasi terdiri dari tiga angka, yakni 1, 2, 3. Misalkan dari populasi yang terdiri dari tiga angka tersebut, akan diambil sampel yang terdiri atas satu angka. Maka sampel-sampel yang mungkin adalah

$$1 \qquad 2 \qquad 3.$$

Diketahui rata-rata dari setiap sampel tersebut adalah

$$1 \qquad 2 \qquad 3.$$

Maka rata-rata dari distribusi sampling rata-rata $\bar{X}$ tersebut adalah

$$\mu_{\bar{X}} = \frac{1+2+3}{3} = 2.$$

Sedangkan standar deviasi dari distribusi sampling rata-rata $\bar{X}$ tersebut adalah

$$\sigma_{\bar{X}} = \sqrt{\frac{(1-2)^2 + (2-2)^2 + (3-2)^2}{3}} = 0{,}8165,$$

yang mana

$$\sigma_{\bar{X}} = \sigma \quad (ketika\, n = 1).$$

Mann dan Lacke (2011:307) menyatakan rumus

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

berlaku ketika paling tidak memenuhi salah satu dari kriteria sebagai berikut.

- ⇨ Jumlah elemen dalam populasi berhingga (*finite*) dan pengambilan elemen untuk sampel dari suatu populasi dengan pengembalian (*with replacement*).

- ⇨ Jumlah elemen dalam populasi tak berhingga (*infinite*) dan pengambilan elemen untuk sampel dari suatu populasi tanpa pengembalian (*without replacement*).

Namun kriteria-kriteria tersebut dapat diganti ketika ukuran sampel kecil (*sample size is small*) dalam perbandingannya terhadap ukuran populasi (*in comparison to the population size*). Ukuran sampel dapat dipandang (*is considered*) kecil dalam perbandingannya terhadap ukuran populasi ketika ukuran sampel lebih kecil atau sama dengan 5% dari ukuran populasi, yakni

$$\frac{n}{N} \leq 0{,}05,$$

dengan n merupakan ukuran sampel dan N ukuran populasi. Namun ketika tidak terpenuhi, maka penghitungan $\sigma_{\bar{X}}$ dihitung dengan rumus

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

di mana

$$\sqrt{\frac{N-n}{N-1}}$$

merupakan faktor koreksi populasi berhingga (Mann dan Lacke, 2011:307).

Berikut diberikan contoh kasus untuk perhitungan standar deviasi dari distribusi sampling $\bar{X}$ dengan rumus $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$. Misalkan suatu populasi terdiri dari tiga angka, yakni 1, 2, 3. Misalkan dari populasi yang terdiri dari tiga angka tersebut, akan diambil sampel yang terdiri atas dua angka dengan pengembalian (*with replacement*). Maka sampel-sampel yang mungkin adalah sebagai berikut.

(1,1)	(1,2)	(1,3)
(2,1)	(2,2)	(2,3)
(3,1)	(3,2)	(3,3)

Perhatikan bahwa karena jumlah elemen dalam populasi berhingga, yakni tiga, dan pengambilan elemen sampel dengan pengembalian, maka standar deviasi dari distribusi sampling rata-rata $\bar{X}$ dihitung dengan rumus sebagai berikut.

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Hasil perhitungan rata-rata untuk setiap sampel sebagai berikut.

1	1,5	2
1,5	2	2,5
2	2,5	3

Maka rata-rata dari distribusi sampling rata-rata $\bar{X}$ tersebut adalah

$$\mu_{\bar{X}} = \frac{1 + 1{,}5 + 2 + 1{,}5 + 2 + 2{,}5 + 2 + 2{,}5 + 3}{9} = \frac{18}{9} = 2.$$

Berikut perhitungan standar deviasi dari distribusi sampling rata-rata $\bar{X}$.

$$\begin{matrix} (1-2)^2 & (1{,}5-2)^2 & (2-2)^2 \\ (1{,}5-2)^2 & (2-2)^2 & (2{,}5-2)^2 \\ (2-2)^2 & (2{,}5-2)^2 & (3-2)^2 \end{matrix}$$

Maka diperoleh hasil sebagai berikut.

1	0,25	0
0,25	0	0,25
0	0,25	1

Sehingga

$$\sigma_{\bar{X}} = \sqrt{\frac{(1-2)^2 + (1{,}5-2)^2 + (2-2)^2 + \cdots + (3-2)^2}{9}}$$

$$\sigma_{\bar{X}} = \sqrt{\frac{1 + 0{,}25 + 0 + 0{,}25 + 0 + 0{,}25 + 0 + 0{,}25 + 1}{9}}$$

$$\sigma_{\bar{X}} = \sqrt{\frac{3}{9}} = \sqrt{0{,}3333333} = 0{,}57735$$

Perhatikan bahwa berdasarkan perhitungan sebelumnya diperoleh

$$\mu_{\bar{X}} = 2$$

$$\sigma_{\bar{X}} = 0{,}57735.$$

Diketahui

$$\mu = \frac{1+2+3}{3} = 2$$

$$\sigma = \sqrt{\frac{(1-2)^2 + (2-2)^2 + (3-2)^2}{3}} = \sqrt{0{,}6666666} = 0{,}81649658.$$

Perhatikan bahwa

$$\sigma_{\bar{X}} \neq \sigma,$$

namun

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

$$0{,}57735 = \frac{0{,}81649658}{\sqrt{2}}$$

$$0{,}57735 = 0{,}57735.$$

Berikut diberikan ilustrasi dalam R.

```
fungsi dasar R.R ×
Source on Save                                   Run
1  library(prob)
2  sampel=urnsamples(c(1,2,3), size=2, replace=TRUE, ordered=TRUE)
3  sampel
4
5  ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
6  ratarata=format(ratarata, digits=4) #pengaturan desimal
7  ratarata
8  mode(ratarata) #perhatikan bahwa tipe data rata-rata adalah character
9  ratarata=as.numeric(ratarata) #mengkonversi tipe data rata-rata, dari character menjadi numeric
10 mode(ratarata)
11 ratarata
12 mean(ratarata) #menghitung rata-rata dari distribusi sampling rata-rata sampel
13 #kode R pada baris 14 untuk menghitung standar deviasi dari distribusi samling rata-rata sampel
14 sd(ratarata)*sqrt((length(ratarata)-1)/(length(ratarata)))
```

Gambar 5.13

```
sampel=urnsamples(c(1,2,3), size=2, replace=TRUE, ordered=TRUE)
sampel
```

```
##   X1 X2
## 1  1  1
## 2  2  1
## 3  3  1
## 4  1  2
## 5  2  2
## 6  3  2
## 7  1  3
## 8  2  3
## 9  3  3
```

```
ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=4) #pengaturan desimal
ratarata
```

```
## [1] "1.0" "1.5" "2.0" "1.5" "2.0" "2.5" "2.0" "2.5" "3.0"
```

```
mode(ratarata) #perhatikan bahwa tipe data rata-rata adalah character
```

```
## [1] "character"
```

```
ratarata=as.numeric(ratarata) #mengkonversi tipe data rata-rata, dari character menjadi numeric
mode(ratarata)
```

```
## [1] "numeric"
```

```
ratarata
```

```
## [1] 1.0 1.5 2.0 1.5 2.0 2.5 2.0 2.5 3.0
```

```
mean(ratarata) #menghitung rata-rata dari distribusi sampling rata-rata sampel
```

```
## [1] 2
```

```
#kode R pada baris 14 untuk menghitung standar deviasi dari distribusi samling rata-rata sampel
sd(ratarata)*sqrt((length(ratarata)-1)/(length(ratarata)))
```

```
## [1] 0.5773503
```

Gambar 5.14

Berikut diberikan contoh kasus untuk perhitungan standar deviasi dari distribusi sampling $\bar{X}$ dengan rumus $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$. Misalkan suatu populasi terdiri dari tiga angka, yakni 1, 2, 3. Misalkan dari populasi yang terdiri dari tiga angka tersebut, akan diambil sampel yang terdiri atas dua angka tanpa pengembalian (*without replacement*). Maka sampel-sampel yang mungkin adalah

(1,2) (1,3) (2,3)

Perhatikan bahwa karena jumlah elemen dalam populasi berhingga, yakni tiga, namun pengambilan elemen sampel tanpa pengembalian, maka standar deviasi dari distribusi sampling rata-rata $\bar{X}$ dihitung dengan rumus sebagai berikut.

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Diketahui rata-rata dari setiap sampel tersebut adalah

1,5 2 2,5,

sehingga rata-rata dari distribusi sampling rata-rata $(\bar{X})$ tersebut adalah

$$\mu_{\bar{X}} = \frac{1{,}5 + 2 + 2{,}5}{3} = \frac{6}{3} = 2.$$

Standar deviasi dari distribusi sampling rata-rata $\bar{X}$ tersebut adalah

$$\sigma_{\bar{X}} = \sqrt{\frac{(1{,}5-2)^2 + (2-2)^2 + (2{,}5-2)^2}{3}}$$

$$\sigma_{\bar{X}} = \sqrt{\frac{0{,}25 + 0 + 0{,}25}{3}}$$

$$\sigma_{\bar{X}} = \sqrt{\frac{0{,}5}{3}} = \sqrt{0{,}16666667} = 0{,}408248.$$

Perhatikan bahwa berdasarkan perhitungan sebelumnya diperoleh

$$\mu_{\bar{X}} = 2$$

$$\sigma_{\bar{X}} = 0{,}408248.$$

Diketahui

$$\mu = \frac{1+2+3}{3} = 2$$

$$\sigma = \sqrt{\frac{(1-2)^2 + (2-2)^2 + (3-2)^2}{3}} = \sqrt{0{,}6666666} = 0{,}81649658.$$

Perhatikan bahwa

$$\sigma_{\bar{X}} \neq \sigma.$$

Namun

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

$$0{,}408248 = \frac{0{,}81649658}{\sqrt{2}} \sqrt{\frac{3-2}{3-1}}$$

$$0{,}408248 = \frac{0{,}81649658}{\sqrt{2}} \sqrt{\frac{1}{2}}$$

$$0{,}408248 = \frac{0{,}81649658}{2}$$

$$0{,}408248 = 0{,}408248$$

Beberapa hal penting mengenai distribusi sampling rata-rata $\bar{X}$, yakni:

- Nilai standar deviasi dari distribusi sampling rata-rata $\bar{X}$ lebih kecil dibandingkan nilai standar deviasi populasi, yakni $\sigma_{\bar{X}} < \sigma$ ketika n lebih besar dari 1. Hal ini terlihat jelas dari rumus

  $$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}.$$

  Sebagai contoh misalkan $\sigma = 20$ dan $n = 4$, maka

  $$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{4}} = 10.$$

  Perhatikan bahwa

  $$\sigma_{\bar{X}} < \sigma$$

  $$10 < 20.$$

- Nilai dari standar deviasi dari distribusi sampling rata-rata $\bar{X}$ akan semakin mengecil ketika ukuran sampel n semakin besar.

  ketika $n \uparrow$ maka $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \downarrow$

Sebagai contoh misalkan $\sigma = 20$ dan $n = 4$, maka

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{4}} = 10.$$

Untuk $n = 20$ maka

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{20}} = 4{,}4721.$$

Untuk $n = 50$ maka

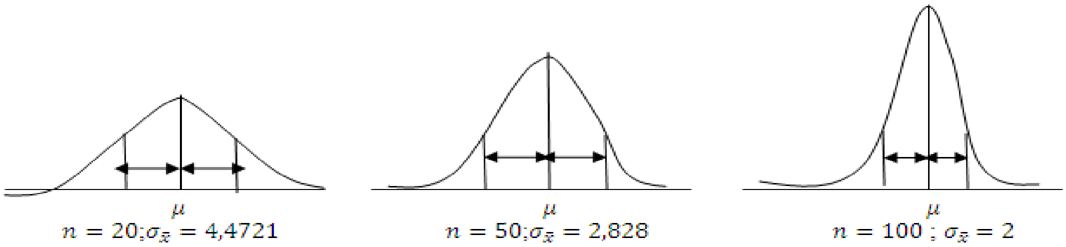
$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{50}} = 2{,}828.$$

Untuk $n = 100$ maka

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{100}} = 2\,.$$

Perhatikan bahwa nilai $\sigma_{\bar{X}}$ semakin mengecil ketika ukuran sampel n semakin besar. Suatu statistik dikatakan estimator konsisten jika nilai standar deviasi dari distribusi sampling statistik tersebut semakin mengecil ketika ukuran sampel n semakin besar, sehingga statistik rata-rata $\bar{X}$ merupakan estimator konsisten dari parameter rata-rata μ (Mann dan Lacke, 2011:307)

Standar deviasi dari distribusi sampling rata-rata $\bar{X}$ merupakan suatu nilai yang mengukur pencaran atau sebaran dari rata-rata sampel dari distribusi sampling rata-rata $\bar{X}$ terhadap rata-rata populasinya μ. Semakin kecil nilai standar deviasi dari distribusi sampling rata-rata $\bar{X}$, maka rata-rata sampel dari distribusi sampling rata-rata $\bar{X}$ semakin mengumpul atau lebih dekat terhadap rata-rata populasinya μ. Pada pembahasan sebelumnya, diketahui untuk untuk $n = 20$ diperoleh $\sigma_{\bar{X}} = 4{,}4721$, untuk $n = 50$ diperoleh $\sigma_{\bar{X}} = 2{,}828$, dan untuk $n = 100$ diperoleh $\sigma_{\bar{X}} = 2$. Perhatikan ilustrasi gambar berikut ini (Gambar 5.15).



Gambar 5.15

Berikut diberikan ilustrasi dalam R (perhatikan Gambar 5.16 dan Gambar 5.17).

```
1  library(prob)
2  sampel=urnsamples(c(1,2,3), size=2, replace=FALSE, ordered=FALSE)
3  sampel
4
5  ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
6  ratarata=format(ratarata, digits=4) #pengaturan desimal
7  ratarata
8  mode(ratarata) #perhatikan bahwa tipe data rata-rata adalah character
9  ratarata=as.numeric(ratarata) #mengkonversi tipe data rata-rata, dari character menjadi numeric
10 mode(ratarata)
11 ratarata
12 mean(ratarata) #menghitung rata-rata dari distribusi sampling rata-rata sampel
13 #kode R pada baris 14 untuk menghitung standar deviasi dari distribusi samling rata-rata sampel
14 sd(ratarata)*sqrt((length(ratarata)-1)/(length(ratarata)))
```

Gambar 5.16

```
sampel=urnsamples(c(1,2,3), size=2, replace=FALSE, ordered=FALSE)
sampel
```

```
##   X1 X2
## 1  1  2
## 2  1  3
## 3  2  3
```

```
ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=4) #pengaturan desimal
ratarata
```

```
## [1] "1.5" "2.0" "2.5"
```

```
mode(ratarata) #perhatikan bahwa tipe data rata-rata adalah character
```

```
## [1] "character"
```

```
ratarata=as.numeric(ratarata) #mengkonversi tipe data rata-rata, dari character menjadi numeric
mode(ratarata)
```

```
## [1] "numeric"
```

```
ratarata
```

```
## [1] 1.5 2.0 2.5
```

```
mean(ratarata) #menghitung rata-rata dari distribusi sampling rata-rata sampel
```

```
## [1] 2
```

```
#kode R pada baris 14 untuk menghitung standar deviasi dari distribusi samling rata-rata sampel
sd(ratarata)*sqrt((length(ratarata)-1)/(length(ratarata)))
```

```
## [1] 0.4082483
```

Gambar 5.17

Bentuk Distribusi Sampling dari Rata-Rata Sampel $\bar{X}$

Mann dan Lacke (2011:310) menyatakan bentuk distribusi sampling dari rata-rata $\bar{X}$ berkenaan (*relates*) atas dua hal, yakni:

⇨ Sampel yang ditarik dari populasi yang berdistribusi normal.

⇨ Sampel yang ditarik dari populasi yang tidak berdistribusi normal.

Jika sampel-sampel yang ditarik berasal dari populasi yang berdistribusi normal dengan rata-rata dan standar deviasi masing-masing μ dan σ, maka:

⇨ Rata-rata distribusi sampling rata-rata $\bar{X}$ sama dengan rata-rata populasi, yakni

$$\mu_{\bar{X}} = \mu.$$

⇨ Standar deviasi distribusi sampling rata-rata $\bar{X}$ sama dengan $\frac{\sigma}{\sqrt{n}}$, dengan asumsi (*assuming*) $n/N \leq 0{,}05$.

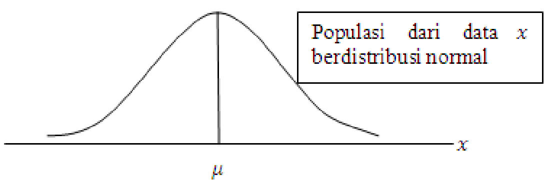
⇨ Bentuk dari distribusi sampling rata-rata $\bar{X}$ berbentuk normal, untuk berapapun ukuran sampel n.

Jadi, jika sampel-sampel yang ditarik berasal dari populasi yang berdistribusi normal dengan rata-rata adalah μ dan standar deviasi adalah σ, maka distribusi sampling dari rata-rata $\bar{X}$ juga terdistribusi secara normal, dengan rata-rata dan standar deviasi

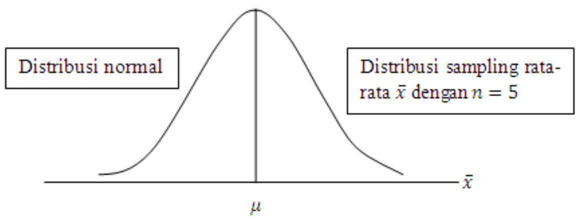
$$\mu_{\bar{X}} = \mu$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad ; \frac{n}{N} \leq 0{,}05.$$

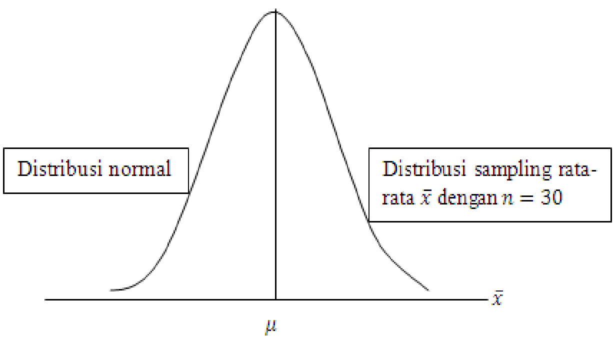
Perhatikan Gambar 5.18 hingga Gambar 5.21.



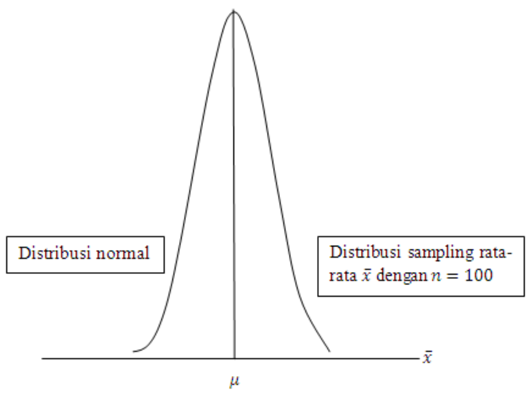
Gambar 5.18



Gambar 5.19



Gambar 5.20



Gambar 5.21

Perhatikan bahwa pada Gambar 5.18 menjelaskan data X berasal dari populasi berdistribusi normal. Pada Gambar 5.19 merupakan kurva dari distribusi sampling rata-rata $\bar{X}$ dengan $n = 5$. Pada Gambar 5.20 merupakan kurva dari distribusi sampling rata-rata $\bar{X}$ dengan $n = 30$.

Pada Gambar 5.21 merupakan kurva dari distribusi sampling rata-rata $\bar{X}$ dengan $n = 100$. Perhatikan bahwa karena sampel-sampel ditarik dari populasi yang berdistribusi normal, maka kurva dari distribusi sampling rata-rata $\bar{X}$ membentuk kurva normal (Gambar 5.19 sampai Gambar 5.21). Perhatikan bahwa standar deviasi dari distribusi sampling rata-rata $\bar{X}$ pada Gambar 5.20 lebih kecil daripada Gambar 5.19, standar deviasi dari distribusi sampling rata-rata $\bar{X}$ pada Gambar 5.21 lebih kecil daripada Gambar 5.20. Perhatikan bahwa semakin besar ukuran sampel, maka akan semakin kecil nilai standar deviasi dari distribusi sampling rata-rata $\bar{X}$. Dalam prakteknya, seringkali populasi yang diteliti tidak berdistribusi normal. Teorema yang sangat penting untuk menyimpulkan bentuk dari distribusi sampling rata-rata $\bar{X}$ adalah **Teorema Limit Sentral** (***Central Limit Theorem***).

Teorema limit sentral menyatakan bahwa untuk sampel berukuran besar, distribusi sampling rata-rata $\bar{X}$ akan mendekati normal, tidak peduli apakah sampel-sampel tersebut ditarik dari populasi yang berdistribusi normal atau tidak, dengan rata-rata dan standar deviasi dari distribusi sampling rata-rata $\bar{X}$ sebagai berikut.

$$\mu_{\bar{X}} = \mu \qquad \text{dan} \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}.$$

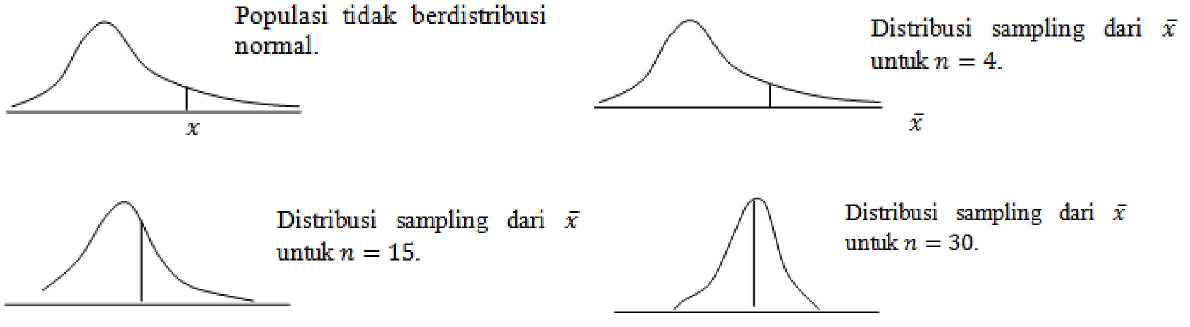
Ukuran sampel n dipertimbangkan cukup besar, yakni $n \geq 30$. Berdasarkan teorema limit sentral, perlu diperhatikan bahwa, jika populasi tidak berdistribusi normal, bentuk dari distribusi sampling rata-rata $\bar{X}$ **tidak secara tepat normal**, **namun mendekati normal**, ketika sampel berukuran besar. Semakin besar ukuran sampel, maka bentuk dari distribusi sampling rata-rata $(\bar{X})$ akan semakin mendekati normal. Berdasarkan teori limit sentral (Mann dan Lacke, 2011:313),

- ⇨ Ketika ukuran sampel $n \geq 30$, maka bentuk dari distribusi sampling rata-rata $(\bar{X})$ mendekati normal, tidak peduli apakah sampel-sampel tersebut ditarik dari populasi berdistribusi normal atau tidak.

- ⇨ Rata-rata dari distribusi sampling rata-rata $(\bar{X})$, yakni $\mu_{\bar{X}}$ sama dengan rata-rata populasi, yakni μ.

- ⇨ Standar deviasi dari distribusi sampling rata-rata $(\bar{X})$, yakni $\sigma_{\bar{X}}$ sama dengan $\sigma/\sqrt{n}$ dengan syarat $n/N \leq 0{,}05$.

Perhatikan ilustrasi gambar berikut.



Gambar 5.22

Berdasarkan Gambar 5.22, populasi tidak berdistribusi normal. Semakin meningkat ukuran sampel, maka distribusi sampling rata-rata $\bar{X}$ semakin berbentuk distribusi normal. Semakin meningkat ukuran sampel, semakin kecil nilai standar deviasi dari distribusi sampling rata-rata $\bar{X}$.

Simulasi Distribusi Sampling dalam R (Bagian 1)

Andaikan diberikan data populasi sebagai berikut.

1,2,3,4,5,6,7,8

Dari data populasi tersebut, **akan diambil sampel yang terdiri dari 2 angka**. Pengambilan sampel **dengan pengembalian** dan **memperhatikan urutan**. Dengan menggunakan R, berikut akan ditentukan seluruh kemungkinan sampel yang mungkin terambil, distribusi frekuensi dari rata-rata sampel, distribusi probabilitas dari rata-rata sampel atau distribusi sampling dari rata-rata sampel, dan disajikan secara visual.

```
fungsi dasar R.R
library(prob)
sampel=urnsamples(c(1,2,3,4,5,6,7,8), size=2, replace=TRUE, ordered=TRUE)
sampel

ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=4) #pengaturan desimal
table(ratarata)
table(ratarata)/length(ratarata)
barplot(table(ratarata))
plot(table(ratarata))
```

Gambar 5.23

```
sampel=urnsamples(c(1,2,3,4,5,6,7,8), size=2, replace=TRUE, ordered=TRUE)
sampel
```

```
##    X1 X2
## 1   1  1
## 2   2  1
## 3   3  1
## 4   4  1
## 5   5  1
## 6   6  1
## 7   7  1
## 8   8  1
## 9   1  2
## 10  2  2
## 11  3  2
## 12  4  2
## 13  5  2
## 14  6  2
## 15  7  2
## 16  8  2
## 17  1  3
## 18  2  3
## 19  3  3
## 20  4  3
## 21  5  3
## 22  6  3
## 23  7  3
## 24  8  3
## 25  1  4
## 26  2  4
## 27  3  4
## 28  4  4
## 29  5  4
## 30  6  4
```

```
## 28  4  4
## 29  5  4
## 30  6  4
## 31  7  4
## 32  8  4
## 33  1  5
## 34  2  5
## 35  3  5
## 36  4  5
## 37  5  5
## 38  6  5
## 39  7  5
## 40  8  5
## 41  1  6
## 42  2  6
## 43  3  6
## 44  4  6
## 45  5  6
## 46  6  6
## 47  7  6
## 48  8  6
## 49  1  7
## 50  2  7
## 51  3  7
## 52  4  7
## 53  5  7
## 54  6  7
## 55  7  7
## 56  8  7
## 57  1  8
## 58  2  8
## 59  3  8
## 60  4  8
## 61  5  8
## 62  6  8
## 63  7  8
## 64  8  8
```

Gambar 5.24

```r
ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=4) #pengaturan desimal
table(ratarata)
```

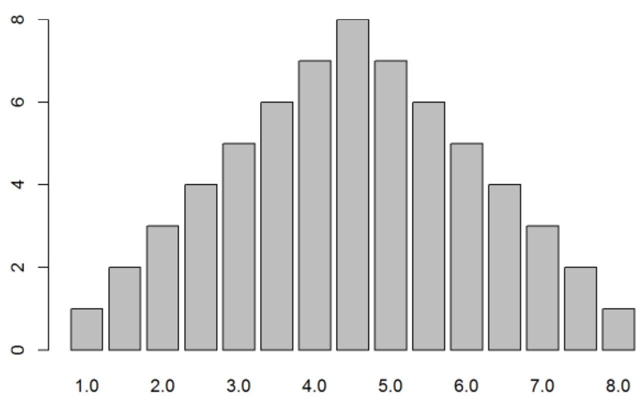
```
## ratarata
## 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0
##   1   2   3   4   5   6   7   8   7   6   5   4   3   2   1
```

```r
table(ratarata)/length(ratarata)
```

```
## ratarata
##      1.0      1.5      2.0      2.5      3.0      3.5      4.0      4.5
## 0.015625 0.031250 0.046875 0.062500 0.078125 0.093750 0.109375 0.125000
##      5.0      5.5      6.0      6.5      7.0      7.5      8.0
## 0.109375 0.093750 0.078125 0.062500 0.046875 0.031250 0.015625
```

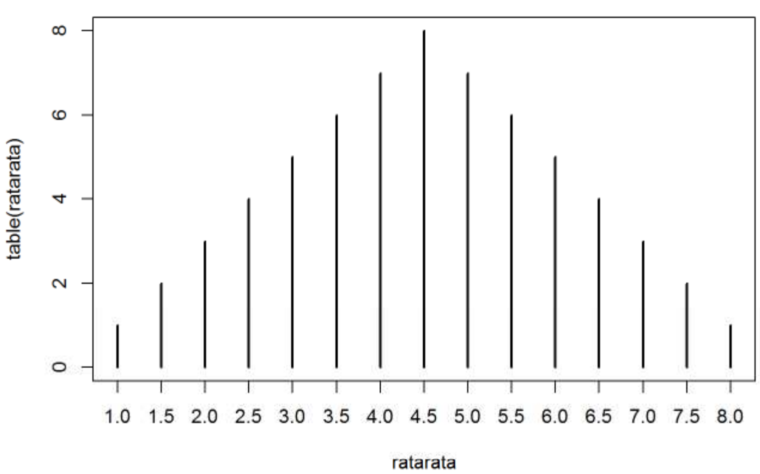
Gambar 5.25

```r
barplot(table(ratarata))
```



Gambar 5.26

```r
plot(table(ratarata))
```



Gambar 5.27

Simulasi Distribusi Sampling dalam R (Bagian 2)

Andaikan diberikan data populasi sebagai berikut.

1,2,3,4,5,6,7,8

Dari data populasi tersebut, **akan diambil sampel yang terdiri dari 3 angka**. Pengambilan sampel **dengan pengembalian** dan **memperhatikan urutan**. Dengan menggunakan R, berikut akan ditentukan seluruh kemungkinan sampel yang mungkin terambil, distribusi frekuensi dari rata-rata sampel, distribusi probabilitas dari rata-rata sampel atau distribusi sampling dari rata-rata sampel, dan disajikan secara visual.

```
library(prob)
sampel=urnsamples(c(1,2,3,4,5,6,7,8), size=3, replace=TRUE, ordered=TRUE)
sampel

ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=4) #pengaturan desimal
table(ratarata)
table(ratarata)/length(ratarata)
barplot(table(ratarata))
plot(table(ratarata))
```

Gambar 5.28

```
sampel=urnsamples(c(1,2,3,4,5,6,7,8), size=3, replace=TRUE, ordered=TRUE)
sampel
```

```
##     X1 X2 X3
## 1    1  1  1
## 2    2  1  1
## 3    3  1  1
## 4    4  1  1
## 5    5  1  1
## 6    6  1  1
## 7    7  1  1
## 8    8  1  1
## 9    1  2  1
## 10   2  2  1
## 11   3  2  1
## 12   4  2  1
## 13   5  2  1
## 14   6  2  1
## 15   7  2  1
## 16   8  2  1
## 17   1  3  1
## 18   2  3  1
## 19   3  3  1
## 20   4  3  1
## 21   5  3  1
## 22   6  3  1
## 23   7  3  1
## 24   8  3  1
## 25   1  4  1
## 26   2  4  1
## 27   3  4  1
## 28   4  4  1
## 29   5  4  1
## 30   6  4  1
## 31   7  4  1
## 32   8  4  1
```

```
## 479   7  4  8
## 480   8  4  8
## 481   1  5  8
## 482   2  5  8
## 483   3  5  8
## 484   4  5  8
## 485   5  5  8
## 486   6  5  8
## 487   7  5  8
## 488   8  5  8
## 489   1  6  8
## 490   2  6  8
## 491   3  6  8
## 492   4  6  8
## 493   5  6  8
## 494   6  6  8
## 495   7  6  8
## 496   8  6  8
## 497   1  7  8
## 498   2  7  8
## 499   3  7  8
## 500   4  7  8
## 501   5  7  8
## 502   6  7  8
## 503   7  7  8
## 504   8  7  8
## 505   1  8  8
## 506   2  8  8
## 507   3  8  8
## 508   4  8  8
## 509   5  8  8
## 510   6  8  8
## 511   7  8  8
## 512   8  8  8
```

Gambar 5.29

```
ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=4) #pengaturan desimal
table(ratarata)
```

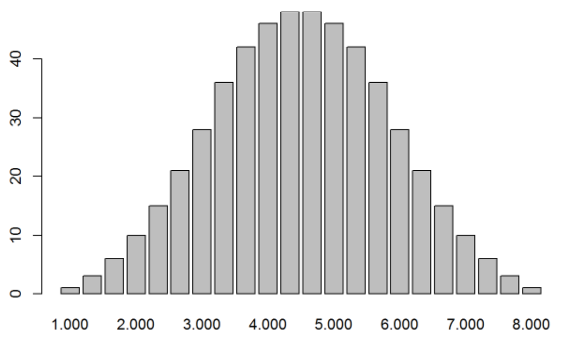
```
## ratarata
## 1.000 1.333 1.667 2.000 2.333 2.667 3.000 3.333 3.667 4.000 4.333 4.667
##     1     3     6    10    15    21    28    36    42    46    48    48
## 5.000 5.333 5.667 6.000 6.333 6.667 7.000 7.333 7.667 8.000
##    46    42    36    28    21    15    10     6     3     1
```

```
table(ratarata)/length(ratarata)
```

```
## ratarata
##       1.000       1.333       1.667       2.000       2.333       2.667
## 0.001953125 0.005859375 0.011718750 0.019531250 0.029296875 0.041015625
##       3.000       3.333       3.667       4.000       4.333       4.667
## 0.054687500 0.070312500 0.082031250 0.089843750 0.093750000 0.093750000
##       5.000       5.333       5.667       6.000       6.333       6.667
## 0.089843750 0.082031250 0.070312500 0.054687500 0.041015625 0.029296875
##       7.000       7.333       7.667       8.000
## 0.019531250 0.011718750 0.005859375 0.001953125
```

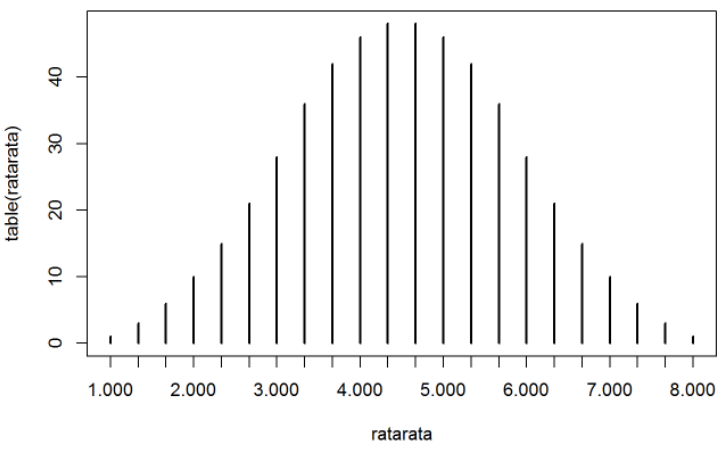
Gambar 5.30

```
barplot(table(ratarata))
```



Gambar 5.31

```
plot(table(ratarata))
```



Gambar 5.32

Simulasi Distribusi Sampling dalam R (Bagian 3)

Andaikan diberikan data populasi sebagai berikut.

1,2,3,4,5,6,7,8

Dari data populasi tersebut, **akan diambil sampel yang terdiri dari 4 angka**. Pengambilan sampel **dengan pengembalian** dan **memperhatikan urutan**. Dengan menggunakan R, berikut akan ditentukan seluruh kemungkinan sampel yang mungkin terambil, distribusi frekuensi dari rata-rata sampel, distribusi probabilitas dari rata-rata sampel atau distribusi sampling dari rata-rata sampel, dan disajikan secara visual.

```
fungsi dasar R.R
Source on Save                                   Run

library(prob)
sampel=urnsamples(c(1,2,3,4,5,6,7,8), size=4, replace=TRUE, ordered=TRUE)
sampel

ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=4) #pengaturan desimal
table(ratarata)
table(ratarata)/length(ratarata)
barplot(table(ratarata))
plot(table(ratarata))
```

Gambar 5.33

```
##      X1 X2 X3 X4
## 1     1  1  1  1
## 2     2  1  1  1
## 3     3  1  1  1
## 4     4  1  1  1
## 5     5  1  1  1
## 6     6  1  1  1
## 7     7  1  1  1
## 8     8  1  1  1
## 9     1  2  1  1
## 10    2  2  1  1
## 11    3  2  1  1
## 12    4  2  1  1
## 13    5  2  1  1
## 14    6  2  1  1
## 15    7  2  1  1
## 16    8  2  1  1
## 17    1  3  1  1
## 18    2  3  1  1
## 19    3  3  1  1
## 20    4  3  1  1
## 21    5  3  1  1
## 22    6  3  1  1
## 23    7  3  1  1
## 24    8  3  1  1
## 25    1  4  1  1
## 26    2  4  1  1
## 27    3  4  1  1
## 28    4  4  1  1
## 29    5  4  1  1
## 30    6  4  1  1
## 31    7  4  1  1
## 32    8  4  1  1
## 33    1  5  1  1
## 34    2  5  1  1
## 35    3  5  1  1
```

```
## 4061  5  4  8  8
## 4062  6  4  8  8
## 4063  7  4  8  8
## 4064  8  4  8  8
## 4065  1  5  8  8
## 4066  2  5  8  8
## 4067  3  5  8  8
## 4068  4  5  8  8
## 4069  5  5  8  8
## 4070  6  5  8  8
## 4071  7  5  8  8
## 4072  8  5  8  8
## 4073  1  6  8  8
## 4074  2  6  8  8
## 4075  3  6  8  8
## 4076  4  6  8  8
## 4077  5  6  8  8
## 4078  6  6  8  8
## 4079  7  6  8  8
## 4080  8  6  8  8
## 4081  1  7  8  8
## 4082  2  7  8  8
## 4083  3  7  8  8
## 4084  4  7  8  8
## 4085  5  7  8  8
## 4086  6  7  8  8
## 4087  7  7  8  8
## 4088  8  7  8  8
## 4089  1  8  8  8
## 4090  2  8  8  8
## 4091  3  8  8  8
## 4092  4  8  8  8
## 4093  5  8  8  8
## 4094  6  8  8  8
## 4095  7  8  8  8
## 4096  8  8  8  8
```

Gambar 5.34

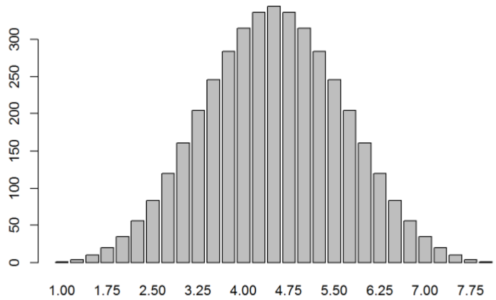
```
ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=4) #pengaturan desimal
table(ratarata)
```

```
## ratarata
## 1.00 1.25 1.50 1.75 2.00 2.25 2.50 2.75 3.00 3.25 3.50 3.75 4.00 4.25 4.50
##    1    4   10   20   35   56   84  120  161  204  246  284  315  336  344
## 4.75 5.00 5.25 5.50 5.75 6.00 6.25 6.50 6.75 7.00 7.25 7.50 7.75 8.00
##  336  315  284  246  204  161  120   84   56   35   20   10    4    1
```

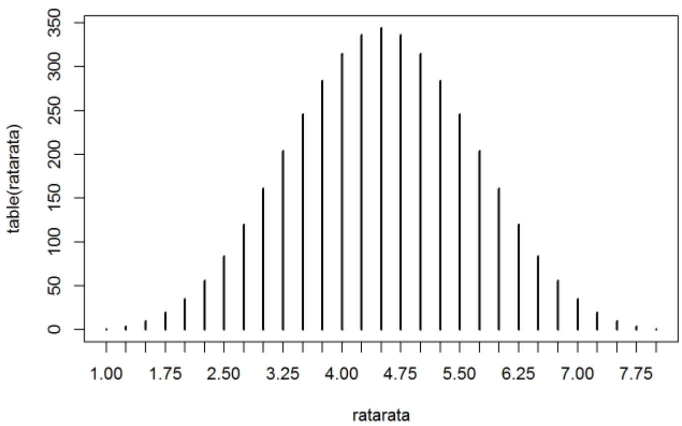
```
table(ratarata)/length(ratarata)
```

```
## ratarata
##         1.00         1.25         1.50         1.75         2.00
## 0.0002441406 0.0009765625 0.0024414062 0.0048828125 0.0085449219
##         2.25         2.50         2.75         3.00         3.25
## 0.0136718750 0.0205078125 0.0292968750 0.0393066406 0.0498046875
##         3.50         3.75         4.00         4.25         4.50
## 0.0600585938 0.0693359375 0.0769042969 0.0820312500 0.0839843750
##         4.75         5.00         5.25         5.50         5.75
## 0.0820312500 0.0769042969 0.0693359375 0.0600585938 0.0498046875
##         6.00         6.25         6.50         6.75         7.00
## 0.0393066406 0.0292968750 0.0205078125 0.0136718750 0.0085449219
##         7.25         7.50         7.75         8.00
## 0.0048828125 0.0024414062 0.0009765625 0.0002441406
```

Gambar 5.35



Gambar 5.36



Gambar 5.37

Simulasi Distribusi Sampling dalam R (Bagian 4)

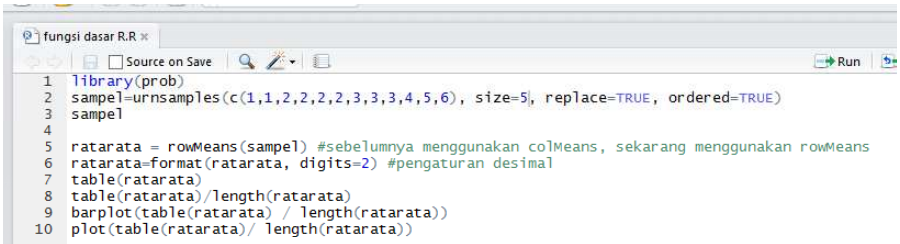
Andaikan diberikan data populasi sebagai berikut.

1,1,2,2,2,2,3,3,3,4,5,6

Dari data populasi tersebut, misalkan:

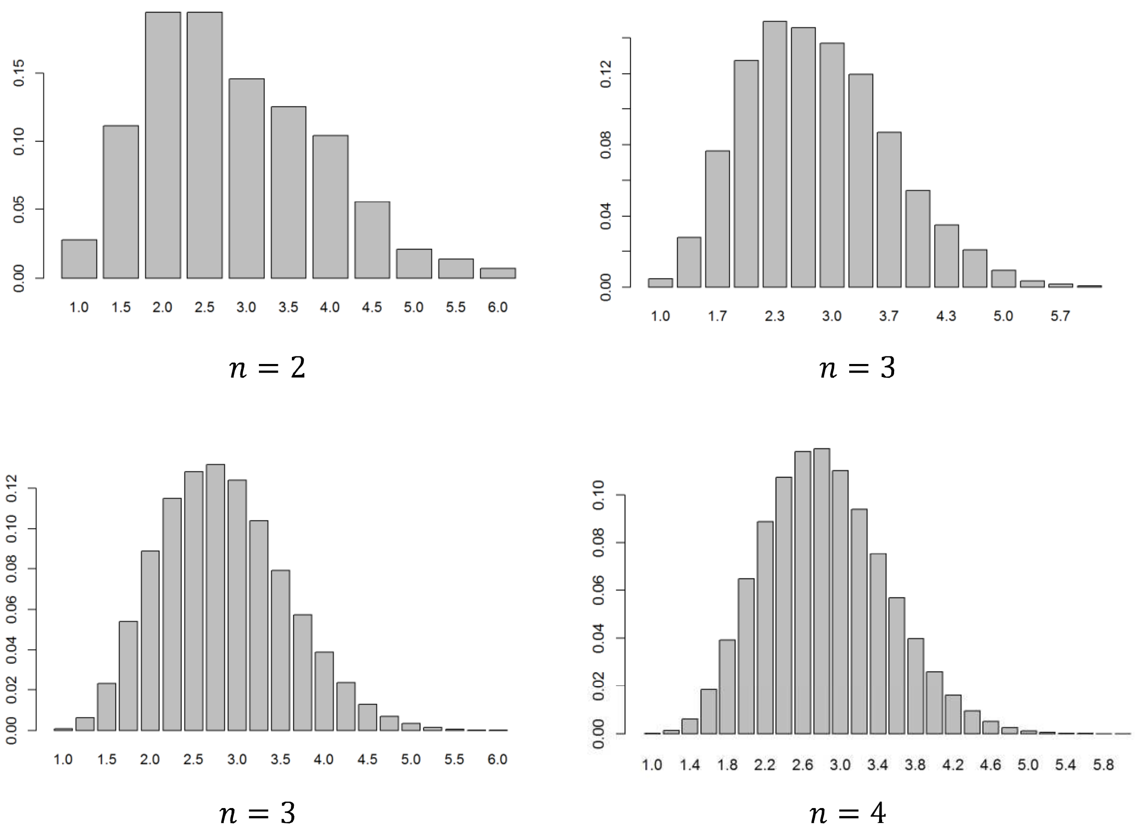
➔**akan diambil sampel yang terdiri dari 2 angka**.
➔**akan diambil sampel yang terdiri dari 3 angka**.
➔**akan diambil sampel yang terdiri dari 4 angka**.
➔**akan diambil sampel yang terdiri dari 5 angka**.

Pengambilan sampel **dengan pengembalian** dan **memperhatikan urutan**. Dengan menggunakan R, berikut akan disajikan secara visual distribusi sampling dari rata-rata sampel.




```
library(prob)
sampel=urnsamples(c(1,1,2,2,2,2,3,3,3,4,5,6), size=5, replace=TRUE, ordered=TRUE)
sampel

ratarata = rowMeans(sampel) #sebelumnya menggunakan colMeans, sekarang menggunakan rowMeans
ratarata=format(ratarata, digits=2) #pengaturan desimal
table(ratarata)
table(ratarata)/length(ratarata)
barplot(table(ratarata) / length(ratarata))
plot(table(ratarata)/ length(ratarata))
```

Gambar 5.38



Gambar 5.39

Referensi

1. Agresti, A. dan B. Finlay. 2009. *Statistical Methods for the Social Sciences, 4th Edition*. United States of America: Prentice Hall.

2. Gio, P.U. dan E. Rosmaini, 2015. Belajar Olah Data dengan SPSS, Minitab, R, Microsoft Excel, EViews, LISREL, AMOS, dan SmartPLS. USUpress.

3. Field, A. 2009. *Discovering Statistics Using SPSS, 3rd Edition*. London: Sage.

4. Johnson, R.A. dan G.K. Bhattacharyya. 2011. *Statistics, Principles and Methods, 6th Edition*. John Wiley and Sons, Inc.

5. Mann, P. S. dan C.J. Lacke. 2011. *Introductory Statistics, International Student Version, 7th Edition*. Asia: John Wiley & Sons, Inc.

6. Montgomery, D. C. dan G. C. Runger. 2011. *Applied Statistics and Probability for Engineers, 5th Edition*. United States of America: John Wiley & Sons, Inc.

7. Ott, R.L. dan M. Longnecker. 2001. *An Introduction to Statistical Methods and Data Analysis*, *5th Edition*. United States of America: Duxbury.

8. Smidth, R. K. dan D. H. Sanders. 2000. *Statistics a First Course, 6th Edition*. United States of America: McGraw-Hill Companies.

9. http://www.dummies.com/how-to/content/how-to-format-numbers-in-r.html

10. http://stackoverflow.com/questions/13033914/sampling-distribution-of-the-sample-mean

11. https://cran.r-project.org/web/packages/prob/prob.pdf