Math 338
Kafai – HW 2

Name
Nhi Vu

Review SAS rand functions and do loops

1. Generate 961 samples of size 625 random numbers from Beta distribution with shape parameters $\alpha = 0.5$, $\beta = 0.5$. For each of these 961 samples calculate the mean.
   a) Find the simulated probability that the mean is between 0.49 and 0.51.
   b) Find the mean of the means.
   c) Find the standard deviation of the means.
   d) Draw the histogram of the means.

2. Generate 961 samples of size 625 random numbers from a Pareto distribution with shape parameter $\alpha = 5$. For each of these 961 samples calculate the mean.
   a) Find the simulated probability that the mean is between 1.25 and 1.26.
   b) Find the simulated mean of the means.
   c) Find the simulated standard deviation of the means.
   d) Draw the histogram of the means.

```
/****** PART 1 ** Nhi Vu *****/

*Generate 961 samples of size 625 random numbers from Beta distribution;
option nodate;
%let ns = 625; /* sample size */
%let n = 961; /* Number of samples*/

data Q1data;
call streaminit(0);
do SampleID = 1 to &n;       /* ID variable for each sample*/
      do i = 1 to &ns;
      x = rand("beta",0.5,0.5);    /*Sample from Beta distribution*/
      output;
      end;
end;
run;


* Calculating mean for each sample s1 to s961;
proc means data= Q1data noprint;
by SampleID;
var x;
output out=OutSummary mean=SampleMean;
run;

proc contents data=OutSummary;
title "Dataset of mean for 961 samples";
run;
```

Dataset of mean for 961 samples

The CONTENTS Procedure

Data Set Name	WORK.OUTSUMMARY	Observations	961
Member Type	DATA	Variables	4
Engine	V9	Indexes	0
Created	03/21/2021 17:10:15	Observation Length	32
Last Modified	03/21/2021 17:10:15	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	WINDOWS_64		
Encoding	wlatin1 Western (Windows)		

Engine/Host Dependent Information

Data Set Page Size	65536
Number of Data Set Pages	1
First Data Page	1
Max Obs per Page	2039
Obs in First Data Page	961
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	C:\Users\921294~1\AppData\Local\Temp\7\SAS Temporary Files_TD5088_AT-CTXXA05_\outsummary.sas7bdat
Release Created	9.0401M6
Host Created	X64_DS12R2
Owner Name	SFSU\921294085
File Size	128KB
File Size (bytes)	131072

Alphabetic List of Variables and Attributes

#	Variable	Type	Len
1	SampleID	Num	8
4	SampleMean	Num	8
3	_FREQ_	Num	8
2	_TYPE_	Num	8

```
/****** PART 1a) ** Nhi Vu *****/

* the simulated probability that the mean is between 0.49 and 0.51;

data a;
set OutSummary (keep=SampleMean);
do i=1 to &n;
     if SampleMean > 0.49 and SampleMean <0.51 then count=1;
     else count=0;
     prob=count/&n;
end;
run;

proc means data=a sum;
title "Question 1.a";
var count prob;
run;
```

Question 1.a

The MEANS Procedure

Variable	Sum
count	491.0000000
prob	0.5109261

```
/****** PART 1b) ** Nhi Vu *****/
* Find the simulated mean of the means;
proc means data=SummaryStat mean;
title "Question 1.b:The simulated mean of the means ";
var SampleMean;
run;
```

Question 1.b:The simulated mean of the means

The MEANS Procedure

Analysis Variable : SampleMean

Mean
0.5002278

```
/****** PART 1c) ** Nhi Vu *****/
* Find the standard deviation of the means;
proc means data=OutSummary std;
title "Question 1.c:The standard deviation of the means";
var SampleMean;
run;
```

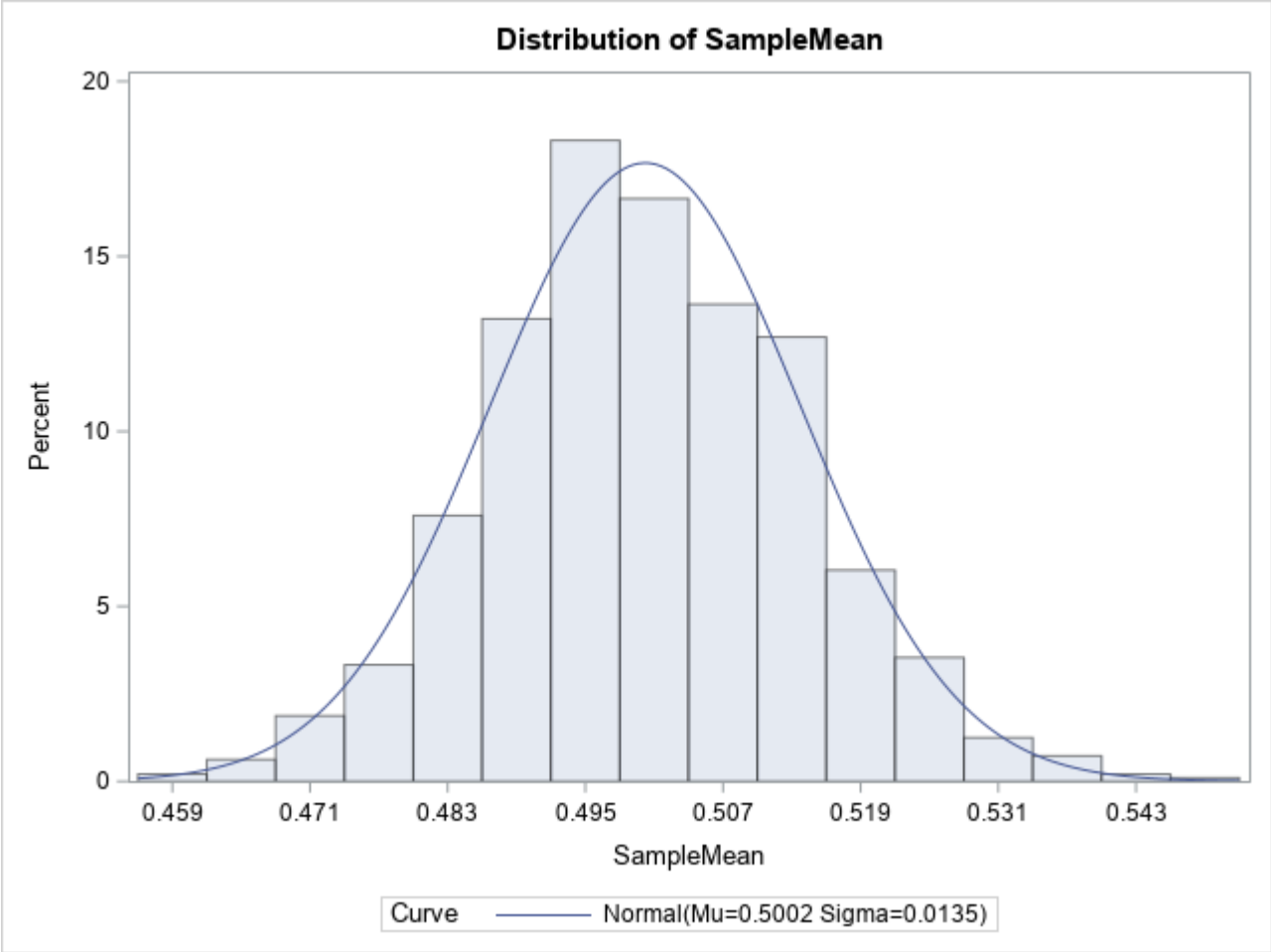
Question 1.c:The standard deviation of the means
The MEANS Procedure

Analysis Variable : SampleMean

Std Dev
0.0135468

```
/****** PART 1d) ** Nhi Vu *****/
* Draw the histogram of the means;

ods select Histogram;
proc univariate data=OutSummary;
title "Question 1.d:Distribution of Sample Mean";
histogram SampleMean / normal ;
run;
quit;
```



```
/****** PART 2 ** Nhi Vu *****/
*Generate 961 samples of size 625 random numbers from a Pareto distribution ;
option nodate;
%let ns = 625; /* sample size */
%let n = 961; /* Number of samples*/

data Q2data;
call streaminit(0);
do SampleID = 1 to &n;      /* ID variable for each sample*/
      do i = 1 to &ns;
      x = rand('pareto',5);   /*Sample from Pareto distribution*/
      output;
      end;
end;
run;

* Calculating mean for each sample;
proc means data=Q2data noprint;
by SampleID;
var x;
output out=SummaryStat mean=SampleMean;
run;
proc contents data=SummaryStat;
title "Dataset of mean for 961 samples";
run;
```

Dataset of mean for 961 samples

The CONTENTS Procedure

Data Set Name	WORK.SUMMARYSTAT	Observations	961
Member Type	DATA	Variables	4
Engine	V9	Indexes	0
Created	03/21/2021 17:32:09	Observation Length	32
Last Modified	03/21/2021 17:32:09	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	WINDOWS_64		
Encoding	wlatin1 Western (Windows)		

Engine/Host Dependent Information

Data Set Page Size	65536
Number of Data Set Pages	1
First Data Page	1
Max Obs per Page	2039
Obs in First Data Page	961
Number of Data Set Repairs	0
ExtendObsCounter	YES
Filename	C:\Users\921294~1\AppData\Local\Temp\7\SAS Temporary Files_TD5088_AT-CTXXA05_\summarystat.sas7bdat
Release Created	9.0401M6
Host Created	X64_DS12R2
Owner Name	SFSU\921294085
File Size	128KB
File Size (bytes)	131072

Alphabetic List of Variables and Attributes

#	Variable	Type	Len
1	SampleID	Num	8
4	SampleMean	Num	8
3	_FREQ_	Num	8
2	_TYPE_	Num	8

```
/****** PART 2a) ** Nhi Vu *****/

* the simulated probability that the mean is between 1.25 and 1.26;

data a;
set SummaryStat (keep=SampleMean);
do i=1 to &n;
     if SampleMean > 1.25 and SampleMean <1.26 then count=1;
     else count=0;
     prob=count/&n;
end;
run;

proc means data=a sum;
title "Question 2.a ";
var count prob;
run;
```

Question 2.a

The MEANS Procedure

Variable	Sum
count	249.0000000
prob	0.2591051

```
/****** PART 2b) ** Nhi Vu *****/
* Find the simulated mean of the means;
proc means data=SummaryStat mean;
title "Question 2.b:The simulated mean of the means ";
var SampleMean;
run;
```

Question 2.b:The simulated mean of the means

The MEANS Procedure

Analysis Variable : SampleMean

Mean
1.2492372

```
/****** PART 2c) ** Nhi Vu *****/
* Find the standard deviation of the means;
proc means data=SummaryStat std;
title "Question 2.c:The standard deviation of the means";
var SampleMean;
run;
```

Question 2.c:The standard deviation of the means

The MEANS Procedure

Analysis Variable : SampleMean

Std Dev
0.0129360

```
/****** PART 2d) ** Nhi Vu *****/
* Draw the histogram of the means;

ods select Histogram;
proc univariate data=SummaryStat;
title "Question 2.d:Distribution of Sample Mean";
histogram SampleMean / normal ;
run;
quit;
```