Week 3: Risk and Evaluation of Alternatives

- Making Decisions in Low-Uncertainty vs. High-Uncertainty Settings
- Example: Evaluating a Wireless Data Plan
- Reward and Risk
- Connecting Random Inputs and Random Outputs
- Simulating Uncertain Outcomes in Excel
- Interpreting Simulation Results: “Short” vs. “Long” Simulations
- Using Histograms to Visualize Simulation Results

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

- See DataPlan10.xlsx

Simulated Data Usage Values and Corresponding Monthly Payment Values (n=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

- We are interested in analyzing the distribution of the **monthly payment**

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

- But first, let us look at the simulated values of **monthly data usage**

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

- Why? Because we know its true probability distribution…

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

- … So we can compare the sample mean and standard deviation with the true values

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)
1	11.9319952
2	24.0282690
3	25.6828047
4	21.7321587
5	34.2335329
6	16.5820597
7	30.7079676
8	36.9010808
9	20.3471859
10	28.3229996
Sample Mean	25.0470054
Sample St. Dev.	7.787935101

- … So we can compare the sample mean and standard deviation with the true values

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)
1	11.9319952
2	24.0282690
3	25.6828047
4	21.7321587
5	34.2335329
6	16.5820597
7	30.7079676
8	36.9010808
9	20.3471859
10	28.3229996
Sample Mean	25.0470054
Sample St. Dev.	7.787935101

In this simulation, **10** values for monthly data usage "drawn" from the normal distribution with mean of 23 and standard deviation of 5 averaged to about **25.047**.

- Sample mean is an approximation to the true value of the expected data usage

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)
1	11.9319952
2	24.0282690
3	25.6828047
4	21.7321587
5	34.2335329
6	16.5820597
7	30.7079676
8	36.9010808
9	20.3471859
10	28.3229996
Sample Mean	25.0470054
Sample St. Dev.	7.787935101

In this simulation, **10** values for monthly data usage "drawn" from the normal distribution with mean of 23 and standard deviation of 5 produced a sample standard deviation of about **7.788**.

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

- The sample mean of the simulated values for monthly payment is about $253, and the sample standard deviation is about $92

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

- The more simulation runs we conduct…

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=10 simulation runs, seed = 123)

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

- … the closer the sample mean and standard deviation will be to the true values

Simulated Data Usage Values and Corresponding Monthly Payment Values (*n*=1000 simulation runs, seed = 123)

	A	B	C	D	E	F
1	**DataPlan_0.xlsx**	**Wireless Data Plan**	**Simulation Run**	**Data Usage, U (GB)**	**Payment, P ($)**	
2	**Operations Analytics MOOC**		1	**11.93199518**	**160**	
3			2	**24.02826903**	**220.4240354**	
4	**Data Allowance (GB)**	20	3	**25.68280473**	**245.242071**	
5	**Fixed Payment ($)**	160	4	**21.7321587**	**185.9823805**	
6	**Rate Above Allowance ($/GB)**	15	5	**34.23353286**	**373.5029929**	
7			6	**16.58205969**	**160**	
8	**Expected Data Usage (GB)**	23	7	**30.7079676**	**320.619514**	
9	**St. Dev. of Data Usage (GB)**	5	8	**36.90108082**	**413.5162123**	
10			9	**20.34718585**	**165.2077878**	
1001			1000	**23.1895728**	**207.843592**	
1002						
1003			**Sample Mean**	**23.28418394**	**220.1594691**	
1004			**Sample St. Dev.**	**4.877547328**	**58.23620041**	

- DataPlan1000.xlsx

Comparing Results for *n*=10 and *n*=1000 simulation runs (seed = 123)

n=10

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

n=1000

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.93199518	160
2	24.02826903	220.4240354
3	25.68280473	245.242071
4	21.7321587	185.9823805
5	34.23353286	373.5029929
6	16.58205969	160
7	30.7079676	320.619514
8	36.90108082	413.5162123
9	20.34718585	165.2077878
1000	23.1895728	207.843592
Sample Mean	23.28418394	220.1594691
Sample St. Dev.	4.877547328	58.23620041

- See DataPlan10.xlsx and DataPlan1000.xlsx

Comparing Results for *n*=10 and *n*=1000 simulation runs (seed = 123)

n=10

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

n=1000

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.93199518	160
2	24.02826903	220.4240354
3	25.68280473	245.242071
4	21.7321587	185.9823805
5	34.23353286	373.5029929
6	16.58205969	160
7	30.7079676	320.619514
8	36.90108082	413.5162123
9	20.34718585	165.2077878
1000	23.1895728	207.843592
Sample Mean	23.28418394	220.1594691
Sample St. Dev.	4.877547328	58.23620041

- The sample mean and sample standard deviation for **monthly data usage** simulated for *n*=1000 runs (approximately, 23.2842 and 4.8775) are much closer to the true values of 23 and 5 than the corresponding sample mean and standard deviation for *n*=10 simulation runs

Comparing Results for *n*=10 and *n*=1000 simulation runs (seed = 123)

n=10

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.9319952	160
2	24.0282690	220.4240354
3	25.6828047	245.242071
4	21.7321587	185.9823805
5	34.2335329	373.5029929
6	16.5820597	160
7	30.7079676	320.619514
8	36.9010808	413.5162123
9	20.3471859	165.2077878
10	28.3229996	284.8449946
Sample Mean	25.0470054	252.9339988
Sample St. Dev.	7.787935101	92.19007977

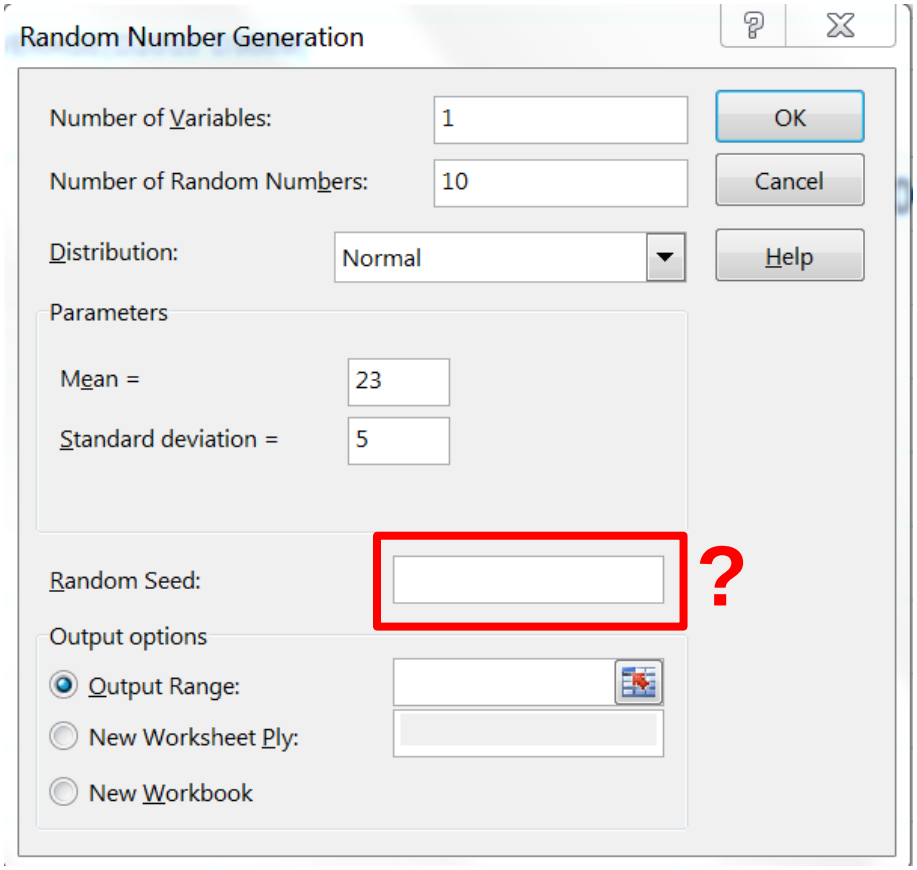
n=1000

Simulation Run	Data Usage, U (GB)	Payment, P ($)
1	11.93199518	160
2	24.02826903	220.4240354
3	25.68280473	245.242071
4	21.7321587	185.9823805
5	34.23353286	373.5029929
6	16.58205969	160
7	30.7079676	320.619514
8	36.90108082	413.5162123
9	20.34718585	165.2077878
1000	23.1895728	207.843592
Sample Mean	23.28418394	220.1594691
Sample St. Dev.	4.877547328	58.23620041

- In a similar way, the sample mean and sample standard deviation for **monthly payment** simulated for *n*=1000 runs (approximately, $220.1995 and $58.2362) are much closer to the true (unknown to us) values than the corresponding sample mean and standard deviation for *n*=10 simulation runs

- Longer simulations produce more precise estimates for the reward and risk measures

Random Seed Value?



- So, what **random seed value** should one use when running a simulation?

Comparing Results for *n*=10 and *n*=1000 simulation runs for different seed values

- Simulated data usage values

***n*=10**	**seed = 123**	**seed = 1826**	**seed = 19104**
Sample Mean, GB	25.05	19.48	24.72
Sample St. Dev., GB	7.79	5.21	3.20

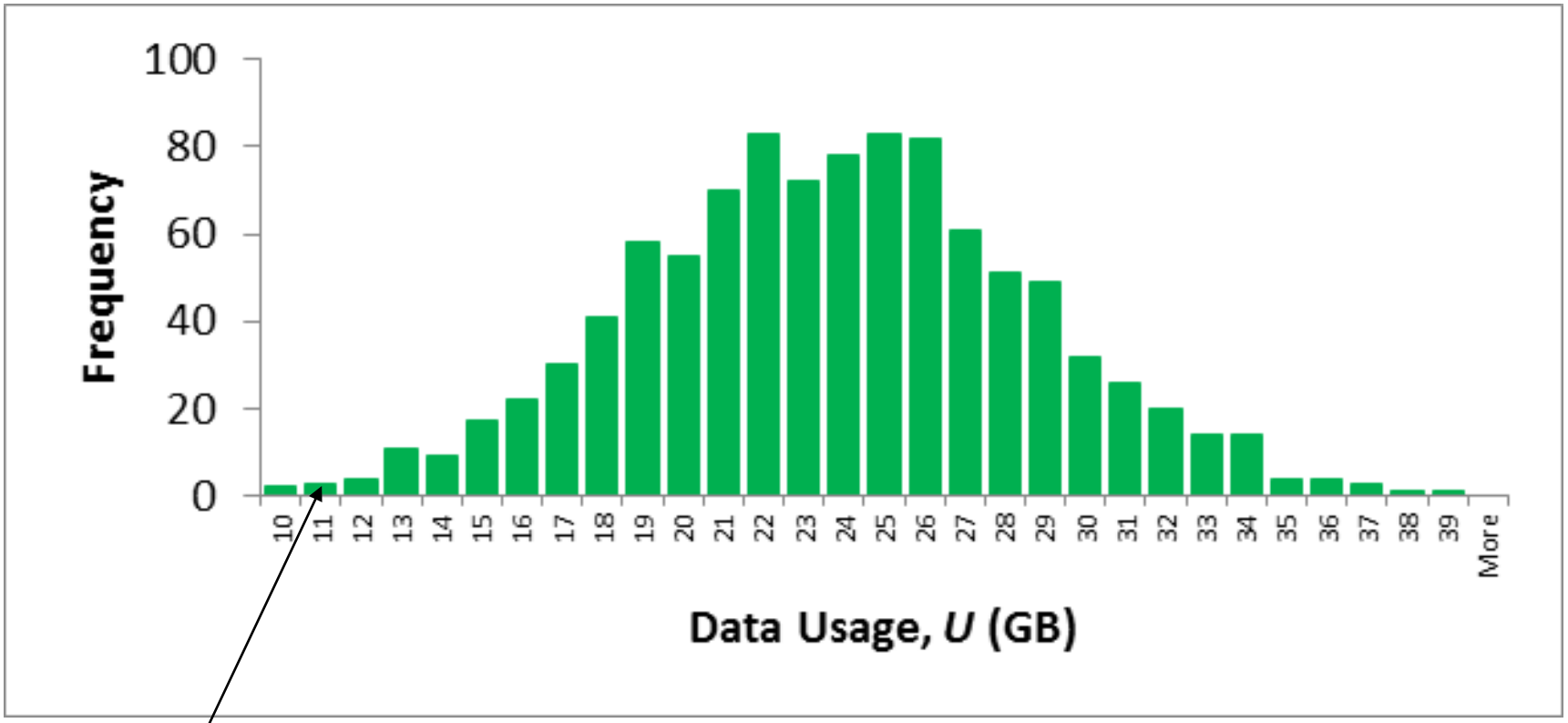
***n*=1000**	**seed = 123**	**seed = 1826**	**seed = 19104**
Sample Mean, GB	23.28	23.08	23.04
Sample St. Dev., GB	4.88	4.90	4.96

- Random seed value does not matter much when you run a simulation with large number of simulation runs

Visualizing Simulation Results Using Histograms

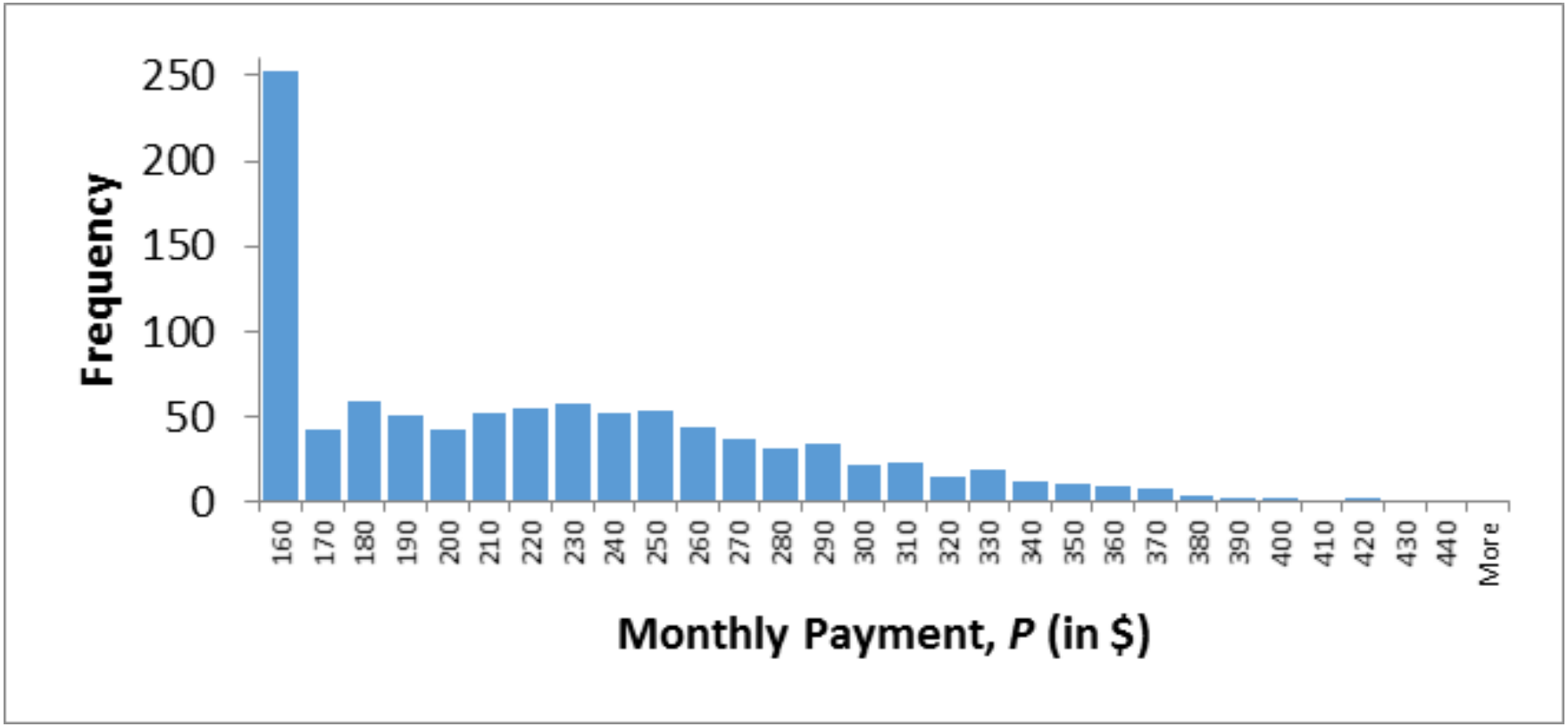
- Histograms are often useful for gaining intuition about the random inputs and the random outputs involved in a simulation
- In the data plan example, the random input is the data usage U, and the random output is the monthly payment P

Histogram of Simulated Values of Data Usage *U* (*n*=1000, seed = 123)



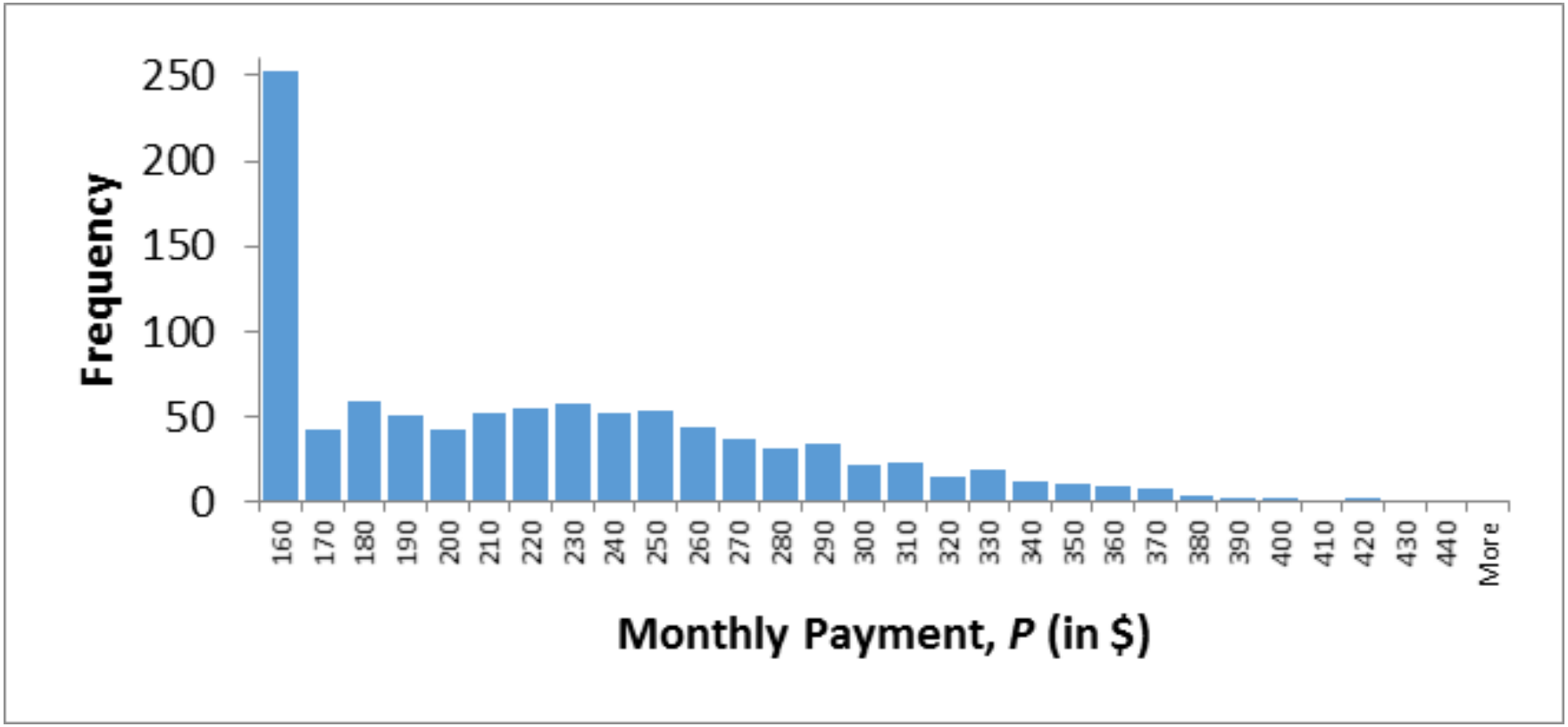
- See DataPlan1000_Histogram.xlsx
- This "bar" indicates the frequency (number of occurrences) for the values of $10 < U \leq 11$

Histogram of Simulated Values of Monthly Payment *P* (*n*=1000, seed = 123)



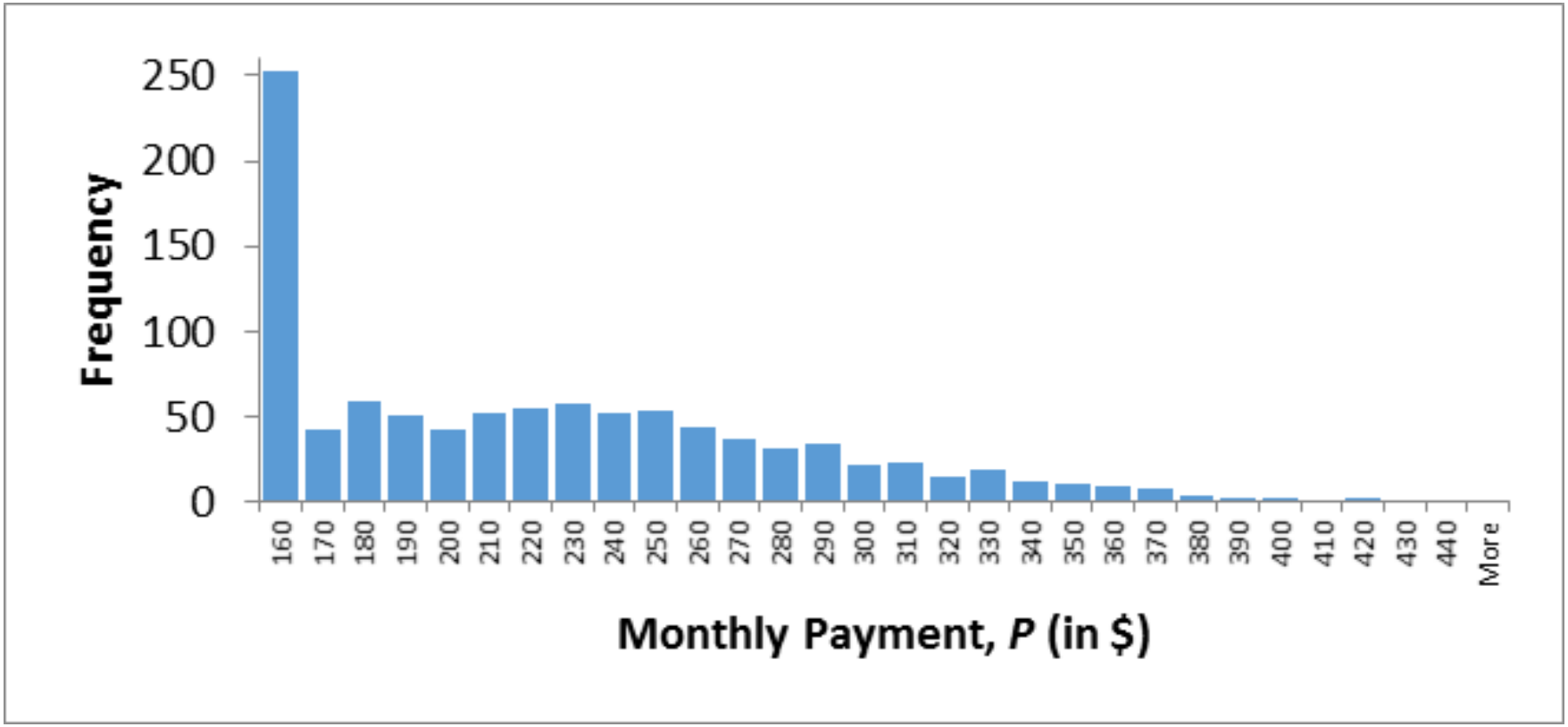
- See DataPlan1000_Histogram.xlsx

Histogram of Simulated Values of Monthly Payment *P* (*n*=1000, seed = 123)



- The input (values of *U*) was drawn from a normal distribution – but the output looks nothing like a normal distribution

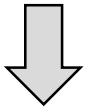
Histogram of Simulated Values of Monthly Payment *P* (*n*=1000, seed = 123)



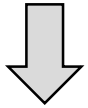
- In general, one must use simulation to understand the shape of the distribution and its parameters for an output random variable

Making Best Decisions in High-Uncertainty Settings: A Roadmap

Decide upon **reward** and **risk** measures



For each competing decision, use **simulation** to estimate reward and risk measures



Use **reward** as an **objective** and **risk measures** as **constraints** to find the best decision

Data Plan Example:

- Reward measure = expected monthly payment
- Risk measure = standard deviation of monthly payment

Data Plan Example:

- Estimate of expected monthly payment = $220.1995
- Estimate of standard deviation of monthly payment = $58.2362

Simulation in Practice: Commercial Simulation Packages

- If you are interested in going beyond Excel's Analysis ToolPak add-in, there exists a number of commercial simulation packages
- A recent comparison of simulation software packages by the *OR/MS Today* can be accessed here*:* http://www.orms-today.org/surveys/Simulation/Simulation.html
- Some commercial packages, in addition to simulating uncertain outcomes, provide capabilities to fit data to a wide range of probability distributions

Simulation in Practice

- Wireless Data Plan example uses one random input (monthly data usage), and analyzes one random output (monthly payment)
- In practice, simulation can be used in the models with many random inputs and many random outputs
- Examples of the use of simulation: articles published in *Interfaces*

Optimizing Capital Investment Decisions at Intel Corporation
Karl G. Kempf, Feryal Erhun, Erik F. Hertzler, Timothy R. Rosenberg, and Chen Peng
Interfaces 2013, 43:1, 62-78.

Kroger Uses Simulation-Optimization to Improve Pharmacy Inventory Management
Xinhui Zhang, Doug Meiser, Yan Liu, Brett Bonner, Lebin Lin
Interfaces 2014, 44:1, 70-84.

- As is the case with optimization, simulation in practice is often used hand-in-hand with other analytics techniques