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 Exce
- Interpreting Simulation Results: "Short" vs. "Long" Simulations
- Using Histograms to Visualize Simulation Results

Simulation Run	Data Usage, U (GB) Payment, P (\$)	Payment, P (\$)
1	11.9319952	160
2	24.0282690	220.4240354
S	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

Simulation Run	Data Usage, U (GB) Payment,	Payment, P (\$)
	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 9229988
Sample St. Dev.	7.787935101	92.19007977

We are interested in analyzing the distribution of the monthly payment

Simulation Run	Data Usage, U (GB) Payment,	Payment, P (\$)
1	11.9319952	160
2	24.0282690	220.4240354
ω	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

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

Why? Because we know its true probability distribution...

Simulation Run	Data Usage, U (GB) Payment,	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

Simulation RunData Usage, U (GB)111.9319952224.0282690325.6828047421.7321587534.2335329616.5820597730.7079676820.34718591028.3229996Sample Mean25.0470054	7.787935101	Sample St. Dev.
nulation Run	25.0470054	Sample Mean
nulation Run		
imulation Run	28.3229996	10
imulation Run	20.3471859	9
imulation Run	36.9010808	8
imulation Run	30.7079676	7
imulation Run	16.5820597	6
imulation Run	33532	5
	21.7321587	4
	25.6828047	ω
	24.0282690	2
	11.9319952	—
	Data Usage, U (GB)	Simulation Run

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

7.787935101	Sample St. Dev.
25.0470054	Sample Mean
28.3229996	10
20.3471859	9
36.9010808	8
30.7079676	7
16.5820597	6
34.2335329	Л
21.7321587	4
25.6828047	ω
24.0282690	2
11.9319952	<u> </u>
Data Usage, U (GB)	Simulation Run

25.047

averaged to about

with mean of 23 and

the normal distribution

standard deviation of 5

usage "drawn" from

values for monthly data

In this simulation, 10

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

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

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.

Simulation Run	Data Usage, U (GB) Payment,	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

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

		•
252.9339988	25.0470054	Sample Mean
284.8449946	28.3229996	10
165.2077878	20.3471859	9
413.5162123	36.9010808	8
320.619514	30.7079676	7
160	16.5820597	6
373.5029929	34.2335329	U
185.9823805	21.7321587	4
245.242071	25.6828047	ω
220.4240354	24.0282690	2
160	11.9319952	Н
ayment, P (\$)	Data Usage, U (GB) Payment, P (\$)	Simulation Run

The more simulation runs we conduct..

Simulation Run	Data Usage, U (GB) Payment,	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

	A	В	С	D	Е	П
\vdash	DataPlan_0.xlsx	Wireless Data Plan Simulation	Run	Data Usage, U (GB) Payment, P (\$)	Payment, P (\$)	
2	Operations Analytics MOOC		1	11.93199518	160	
ω			2	24.02826903	220.4240354	
4	Data Allowance (GB)	20	ω	25.68280473	245.242071	
5	Fixed Payment (\$)	160	4	21.7321587	185.9823805	
6	Rate Above Allowance (\$/GB) 15	15	5	34.23353286	373.5029929	
7			6	16.58205969	160	
∞	Expected Data Usage (GB)	23	7	30.7079676	320.619514	
9	St. Dev. of Data Usage (GB)	ъ	∞	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

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

92.19007977	7.787935101	Sample St. Dev.
252.9339988	25.0470054	Sample Mean
284.8449946	28.3229996	10
165.2077878	20.3471859	9
413.5162123	36.9010808	000
320.619514	30.7079676	7
160	16.5820597	0 = 10
373.5029929	34.2335329	5
185.9823805	21.7321587	4
245.242071	25.6828047	3
220.4240354	24.0282690	2
160	11.9319952	1
Payment, P (\$)	Data Usage, U (GB) Payment, P (\$)	Simulation Run D

Simulation Run	Simulation Run Data Usage, U (GB) Payment, P (\$)	Payment, P (\$)
1	11.93199518	160
2	24.02826903	220.4240354
ω	25.68280473	245.242071
4	21.7321587	185.9823805
5 n=1000	00 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

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

	4	
252.9339988	25.0470054	Sample Mean
284.8449946	28.3229996	10
165.2077878	20.3471859	9
413.5162123	36.9010808	8
320.619514	30.7079676	7
160	16.5820597	$_{6}$ $n=10$
373.5029929	34.2335329	5
185.9823805	21.7321587	4
245.242071	25.6828047	ω
220.4240354	24.0282690	2
160	11.9319952	P
Payment, P (২)	Data Usage, U (GB) Payment, P (\$)	Simulation Run

58.23620041	4.877547328	Sample St. Dev.
220.1594691	23.28418394	Sample Mean
207.843592	23.1895728	1000
165.2077878	20.34718585	9
413.5162123	36.90108082	8
320.619514	30.7079676	7
160	16.58205969	6
373.5029929	00 34.23353286	5 n=1000
185.9823805	21.7321587	4
245.242071	25.68280473	ω
220.4240354	24.02826903	2
160	11.93199518	1
Payment, P (\$)	Data Usage, U (GB) Payment, P (\$)	Simulation Run

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

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

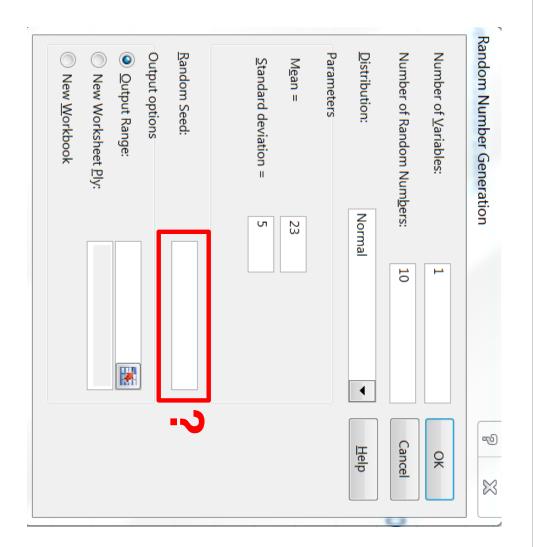
92.19007977	7.787935101	Sample St. Dev.
252.9339988	25.0470054	Sample Mean
284.8449946	28.3229996	10
165.2077878	20.3471859	
413.5162123	36.9010808	
320.619514	30.7079676	
160	16.5820597	<i>n</i> =10
373.5029929	34.2335329	
185.9823805	21.7321587	
245.242071	25.6828047	
220.4240354	24.0282690	
160	11.9319952	
Payment, P (\$)	Data Usage, U (GB) Payment, P (\$)	Simulation Run Da

Cimulation Dun		Daymont D (¢)
	Silidiation van Data Osage, O (OD) i ayılıcılı, i (7)	rayment, r (2)
1	11.93199518	160
2	24.02826903	220.4240354
3	25.68280473	245.242071
4	21.7321587	185.9823805
5 n=1000	00 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

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

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?

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

Simulated data usage values

<i>n</i> =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

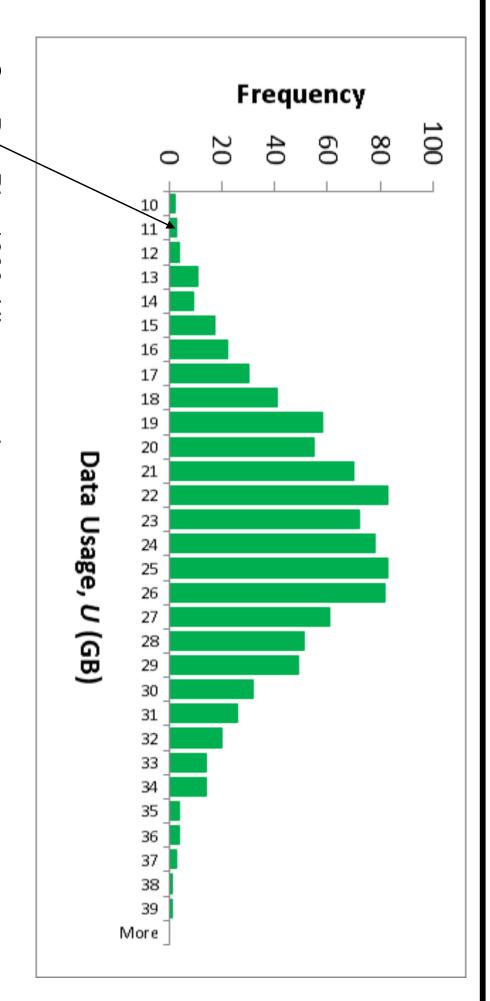
<i>n</i> =1000	seed = 123	seed = 1826	seed = 19104
Sample Mean, GB	23.28	23.08	23.04
Sample St. Dev., GB	4.88	4.90	4.96

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

Visualizing Simulation Results Using Histograms

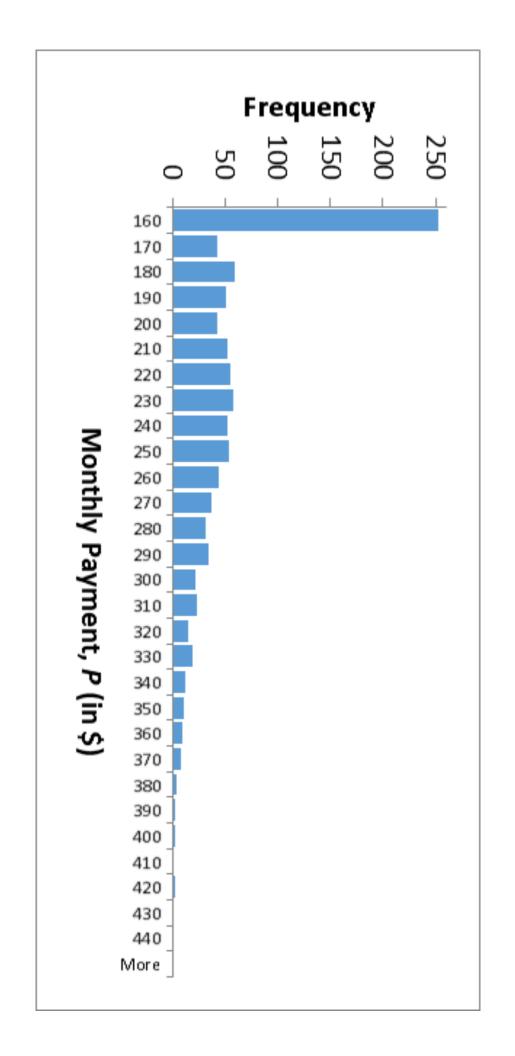
- Histograms are often useful for gaining intuition about the random inputs and the random outputs involved in a simulation
- output is the monthly payment P In the data plan example, the random input is the data usage *U*, and the random

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



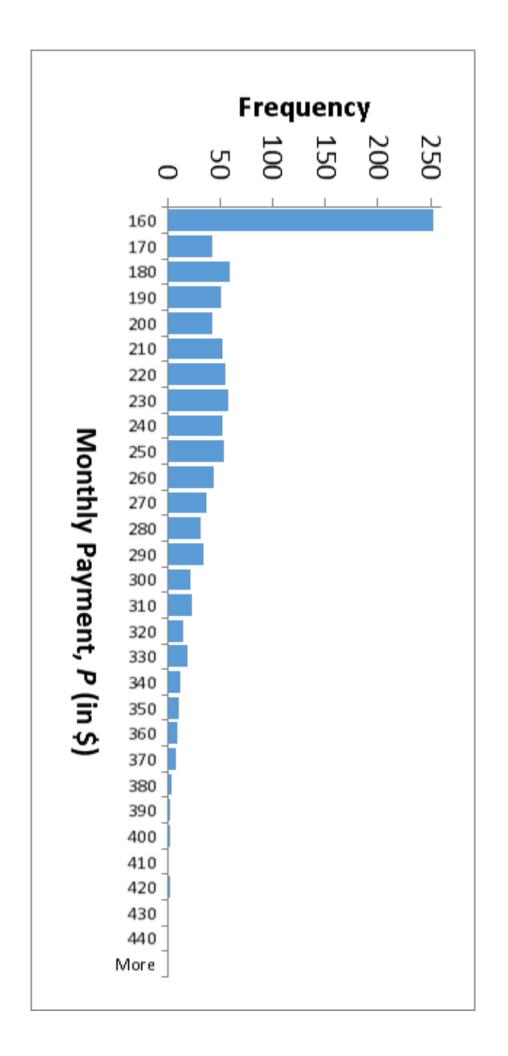
- ♦ See DátaPlan1000_Histogram.xlsx
- 10 < *U* ≤ 11 This "bar" indicates the frequency (number of occurrences) for the values of

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



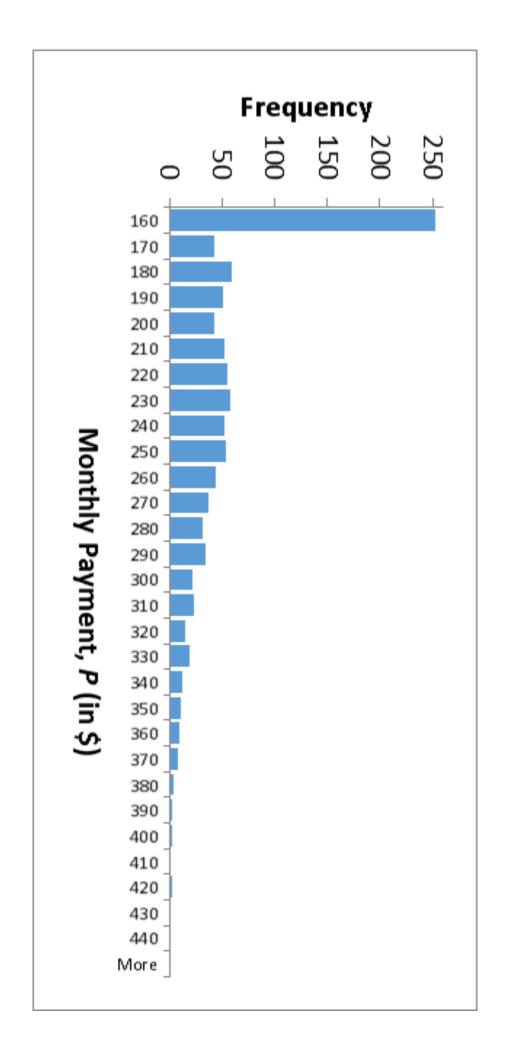
See DataPlan1000_Histogram.xlsx

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)



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

Settings: A Roadmap Making Best Decisions in High-Uncertainty

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

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

Simulation in Practice

- analyzes one random output (monthly payment) Wireless Data Plan example uses one random input (monthly data usage), and
- 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