

Uncertainty and Sensitivity Analyses

Marcelo Coca Perrailon

University of Colorado
Anschutz Medical Campus

Cost-Effectiveness Analysis

HSMP 6609

2020

Outline

- Defining terms: variability, heterogeneity, uncertainty
- Sensitivity analysis: “deterministic” and “probabilistic”
- Base case, one-way, two-way, three-way, scenarios
- Influential variables: tornado diagrams
- More advanced methods: probabilistic sensitivity analysis (PSA)
- Probability distributions
- How to choose probability distributions for probabilistic sensitivity analysis
- Example: HIV model

What is sensitivity analysis?

- We have already done sensitivity analyses in homework and during class
- It is essentially the study of **how changes in model inputs affect model outputs**
- **Inputs** are quantities, prices, probabilities, and so on. In essence, all we measure in a CEA study
- **Outputs** are the model *outcomes* we care about: average or expected costs, average or expected years of life gained, average or expected ICER or Net Monetary Benefit. Most often than not, the **focus is on ICER**
- The objective of a sensitivity analysis is to understand 1) which are the **most important inputs/parameters** that affect our study/model and 2) quantify the **overall uncertainty** of our models
- We often **cannot do traditional statistical analyses** (p-values, confidence intervals, etc) because we don't have individual data

Deterministic and probabilistic sensitivity analysis

- We'll cover two types of sensitivity analyses:
 - 1 **Deterministic**: We choose values for one or more parameters keeping the rest constant. For example, min or max or a case that has policy relevance. This is what we have done so far
 - 2 **Probabilistic**: We assign parameters a probability distribution and use simulations to compute new ICERs or other outcomes of interest
- One confusing part is that, **once again**, we are dealing with concepts that are called different names in different fields
- We need to spend some time defining terms

Defining terms

- **Variability:** Variation due to chance; random variability. Even with identical patients facing the same probability of an event, some experience the event and some do not. Sometimes called *stochastic uncertainty* or first-order uncertainty (less common)
- **Heterogeneity:** Differences between patients that can be attributed or explained by patient's characteristics (think sex, age, income, and so on)
- **Uncertainty:** What sensitivity analysis tries to measure. We **do not know the true value** of some input (parameter) or the true way a process is generated
- Two types of uncertainty: **parameter uncertainty** and **model uncertainty**

Parameter uncertainty

- We do not know the true value of a parameter. For example, the true average cost for a procedure in the population
- We could **estimate** the mean cost in a sample but the estimated parameter (the mean) itself has some variability
- In statistics, a measure of **parameter** uncertainty is the **standard error** (do not confuse it with the standard deviation)
- We can reduce some of this uncertainty if we had a larger sample; the standard error depends on sample sizes. The larger the sample the more precise the estimate and the smaller the standard error
- (A related concept: **consistency**. In probability, the larger the sample size, the closer the estimated parameter will be to its true value)
- **For the purpose of this class:** Each model parameter has some uncertainty that we somehow **should take into account**

Parameter uncertainty

- Last week when we covered the SIR/SEIR models used to understand the coronavirus, I mentioned **parameter uncertainty**
- We just don't have good information right now (as of mid April) to build good models
- What is the transmission rate, β ? What is the recovery, γ ? Even more basic, what is the mortality rate?
- In the HIV example: is the relative risk really 0.509? What was the confidence interval of the study? Intuitively, if the 95% confidence interval was, say, (0.45 - 0.55) we'd argue that the parameter uncertainty less than if it were (0.2 - 0.70)
- All that is parameter uncertainty
- (I'm going to say more about confidence intervals in about 30 mins)

Model uncertainty

- **Models are simplifications of reality**; some models capture the key elements of reality and some don't
- Any model that you see in the literature could be modified to make it more “realistic” but it adds complexity. More realistic models are not necessarily better
- The uncertainty we have about *model assumptions* is called **model uncertainty**
- For example, when we were modeling breast cancer recurrence, we were wondering if not keeping track of time from recurrence was a problem—we used tunnels to sidestep that problem. Was it the wrong assumption?
- Is ignoring the Exposed health state in the SIR model a terrible mistake?

Analogy with regression

- Mandatory slide in cost-effectiveness... **Cover your ears if you don't know regression analysis** (but don't get too distracted)

- A standard linear regression model can be written as:

$$y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \epsilon_i$$

where ϵ_i is assumed to distribute normal with mean 0 and variance σ^2

- After collecting data on variables y and X_1 to X_k , the estimated model is:

- $E[\hat{y}] = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_k X_k$

- Standard errors for the coefficients $SE(\hat{\beta}_j)$ are also estimated

- Each coefficient is a random variable with a (theoretically) known distribution. Once you know the estimated coefficients and their SEs, it is possible to test hypotheses, build confidence intervals, etc

Analogy with regression

- The random error ϵ corresponds to *stochastic uncertainty* or variability due to chance
- *Parameter uncertainty* is measured by the standard error $SE(\hat{\beta}_j)$ of the estimated parameter $\hat{\beta}_j$
- *Heterogeneity* of effects is represented by coefficients $\hat{\beta}_j$
- *Model uncertainty* is the uncertainty around the assumptions of the model. Is it correct to assume that the error, and consequently the outcome Y , distribute normal? Are we including all the relevant covariates X ? Should we model Y or $\log(Y)$?
- Anyhow, most textbooks and articles introducing uncertainty use this analogy so there you go

Cheat sheet

- Of course, there is an article trying to come up with an agreement on names
- From Briggs et al (2012). I used red to underline what I think is the most common term (apparently I don't agree with Briggs)

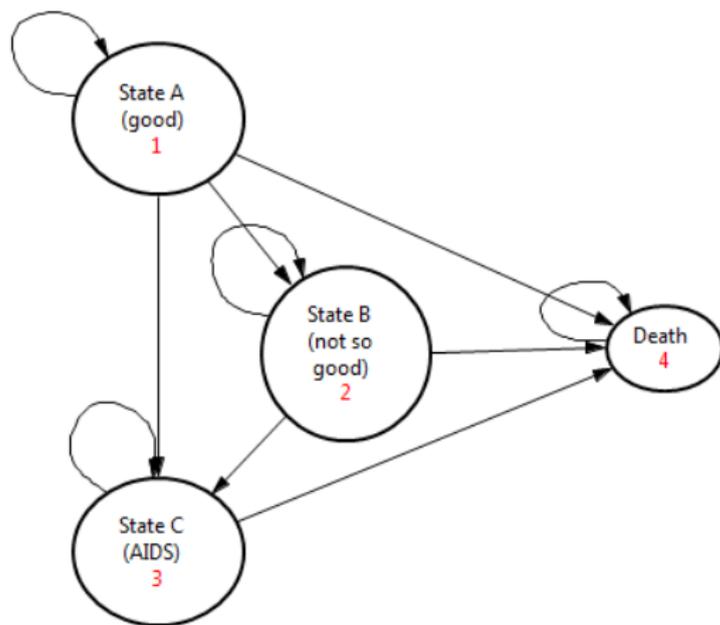
Table 1 – Uncertainty for decision modeling: Concepts and terminology.

Preferred term	Concept	Other terms sometimes employed	Analogous concept in regression
Stochastic uncertainty	<u>Random variability in outcomes between identical patients</u>	Variability <u>Randomness</u> Monte Carlo error First-order uncertainty	Error term
<u>Parameter uncertainty</u>	The uncertainty in estimation of the parameter of interest	Second-order uncertainty	Standard error of the estimate
<u>Heterogeneity</u>	The variability between patients that can be attributed to characteristics of those patients	Variability Observed or explained heterogeneity	Beta coefficients (or the extent to which the dependent variable varies by patient characteristics)
Structural uncertainty	The assumptions inherent in the decision model	<u>Model uncertainty</u>	The form of the regression model (e.g., linear, log-linear)

Deterministic sensitivity analysis

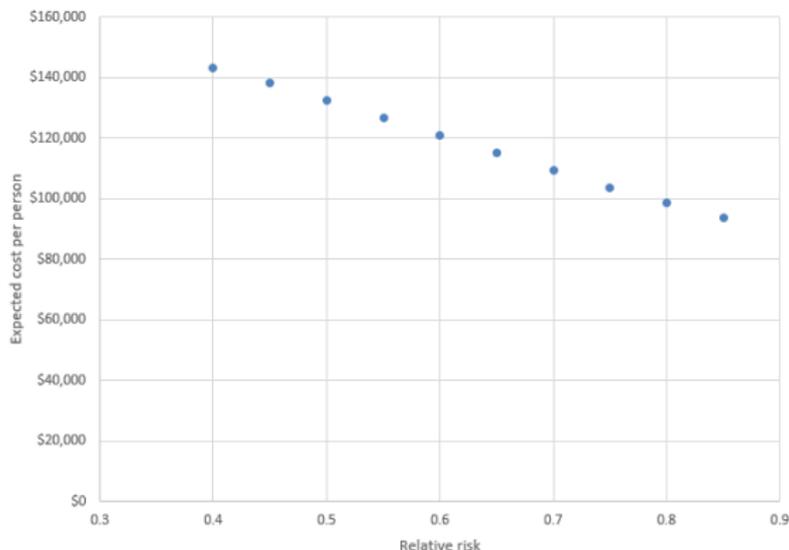
- We'll focus on simple deterministic sensitivity analyses
- **One-way**: Change one parameter at a time keeping all others constant
- Standard way of presenting one-way sensitivity analyses results is to plot the parameter you are changing in the x-axis and an output of interest on the y-axis
- In the HIV example, we could change the relative risk and analyze what happens to average costs in the combination therapy group and the ICER
- You did that in one homework

HIV transition diagram, reminder



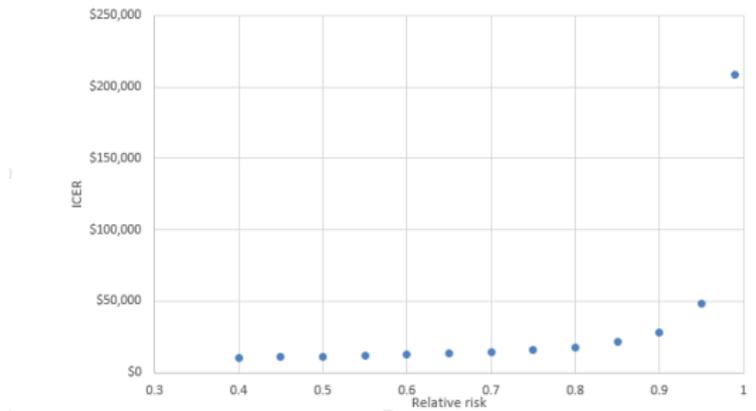
- See Excel file for more details

One-way sensitivity analysis of relative risk with respect combination therapy **cost**



- Lower RR implies combo therapy more effective, so more people stay in states that are costly. It would be unrealistic to choose very high or very low relative risks. Combination therapy at the time of the trial was not a miracle medication

One-way sensitivity analysis of relative risk with respect to ICER



- ICER shows more nuance because of more moving parts. More effective drug increases costs but it also increases life expectancy. Less effective combo therapy becomes more similar to mono therapy but costs keep going up

For special assignment II

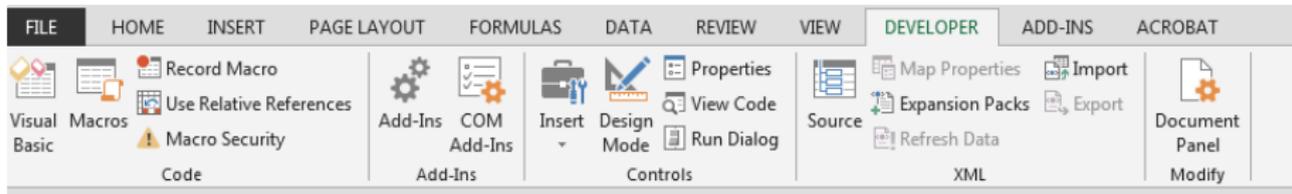
- Most of the time we want to understand how changing a parameter changes the ICER, not just one component of the ICER
- The ICER has four components: $\frac{C_{combo} - C_{mono}}{Years_{combo} - Years_{mono}}$
- You need to understand how the four components are changing
- Some analyses result in weird-looking graphs and non-linearity, usually because of dividing by number close to zero
- A ratio in which the denominator is close to zero tends to infinity, positive or negative
- In the HIV example, if the relative risk is close to 1, then there is no difference in effectiveness between mono and combo therapy. That means that $Years_{combo} - Years_{mono} \sim 0$

Using a macro for sensitivity analysis in Excel

- It is easy to change values in Excel but it is **not the most practical way** to perform sensitivity analyses
- Excel has a **macro language** that is very flexible and relatively easy to use, although it can take a while to master it
- It's based on a version of a programming language, Visual Basic (Visual Basic for Applications, VBA)
- Fortunately, you don't need to master it in order to use it for simple analyses
- We will use a simple macro for quick sensitivity analyses
- But it's worth learning macros. **Add it to resume** (if you learned how to use macros, of course)

Macros in Excel

- First, make sure that you can see the Developer ribbon in Excel:
 - 1 Click the File tab.
 - 2 Click Options.
 - 3 Click Customize Ribbon.
 - 4 Under Customize the Ribbon and under Main Tabs, select the Developer check box



Record a macro

- The easiest way to use macros is to **record** a macro of a **repetitive task** like changing the RR and then copying and pasting the resulting ICER into a new column
- Excel will create VBA code to do the repetitive task
- It is then (relatively) easy to modify the code so the repetitive task is, well, repeated several times
- See the Excel file HIV “*sensitivity_m.xlsm*” for an example (macros could be used as a computer virus so you’ll get a warning before opening the file)
- See the video “Recording a macro” to see how I created the macro (one good thing about the coronavirus: I learned how to make screen videos)

Final code

```
Sub oneway()  
,  
' oneway Macro  
' Performs a basic sensitivity analysis  
,  
' Keyboard Shortcut: Ctrl+Shift+O  
  
Dim i As Integer  
  
For i = 5 To 17  
    Range("D" & i).Select  
    Selection.Copy  
    Sheets("Markov").Select  
    Range("D26").Select  
    ActiveSheet.Paste  
    Range("U32").Select  
    Application.CutCopyMode = False  
    Selection.Copy  
    Sheets("Macro").Select  
    Range("E" & i).Select  
    Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _  
        :=False, Transpose:=False  
Next i  
  
End Sub
```

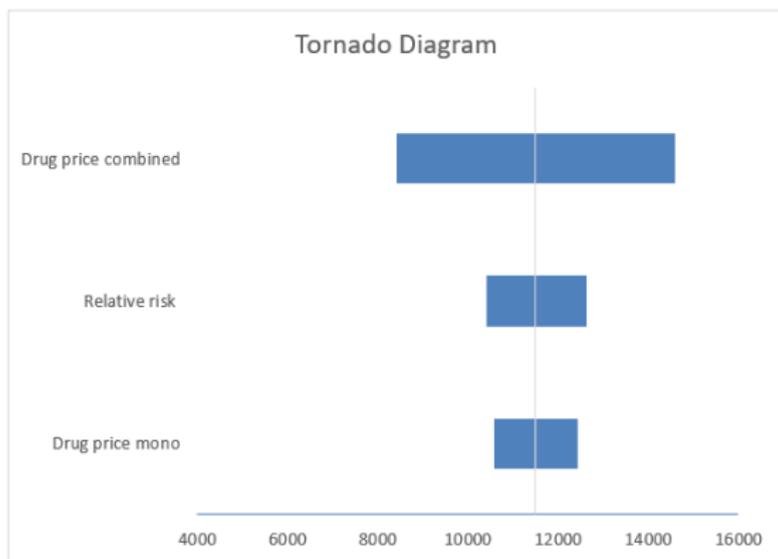
Logic

- For each copy-paste step, Excel creates a chunk of code
- We used this chunk of code to create a loop so Excel does the same in a range of cells, from rows 5 to 17
- We added
 - 1 Dim i As Integer
 - 2 For i = 5 To 17
 - 3 Next i
- See <http://www.excel-easy.com/vba.html> for more Examples
- Remember, this code only repeats interactions with Excel that you could do “by hand” (pointing, clicking, or using the keyboard)
- VBA is quite flexible so you can do many more things

Tornado diagrams

- Tornado diagrams are a nice way of depicting results of one-way sensitivity analyses
- Provides a way to quickly figure out which **parameters have the most influence** on outcomes
- This is very helpful because one needs to pay more attention to the most important parameters
- In other words, we may not need to worry too much about one parameter value if it turns out that the parameter is not that “important”
- We will do a tornado diagram of the the HIV model

Finished tornado diagram



- The top variable has more uncertainty or is more sensitive to possible range of values

Tornado diagram HIV model

- We need to come up with a **range of values** for the parameters
- To make things simpler, I'll focus on only three variables
- I came up with a range for them:

Variable	Base case	Lower-bound	Upper-bound
Drug price mono	2278	1595	2961
Drug price combined	4364	3055	5673
Relative risk	0.509	0.4	0.6

Where do we get these numbers?

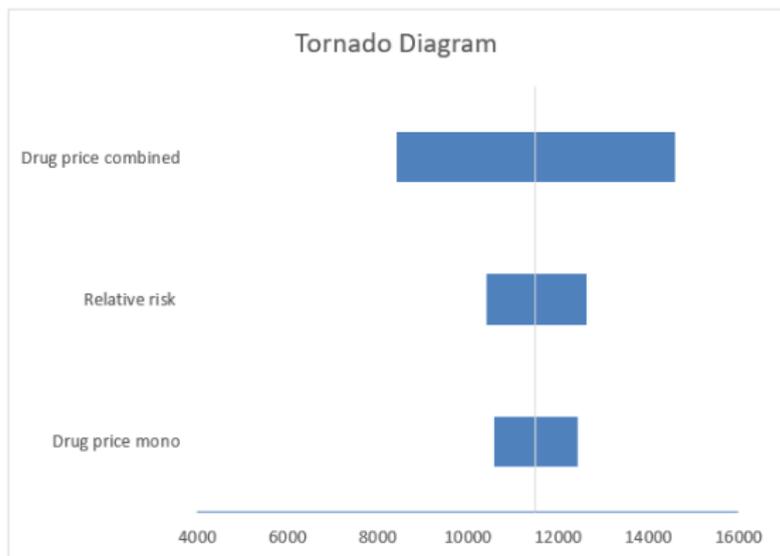
- The **base-case** scenario is our best estimate (sometimes called **baseline**) scenario or baseline values
- The range could be obtained from multiple sources, like drug prices that are common for hospitals or prices after discounts
- The relative risk was obtained from a meta-analysis, so a **confidence interval could be available**
- We can then calculate the ICER corresponding to each of the values in the table to obtain a lower and upper bound estimate of the new ICER corresponding to lower and upper bound range of values

Table for tornado diagram

Variable	ICER-Lb	ICER-Ub	Range
Drug price mono	12426	10572	1854
Relative risk	10422	12621	2199
Drug price combined	8413	14585	6172

- Lb = lower bound; Ub = upper bound
- The range is the absolute value of the difference between the upper and the lower bound
- Table is sorted from lowest to largest **range**
- **A tornado diagram is a graphical representation of this table**

Tornado diagram



- Vertical line is the ICER for base case (\$11,499); the most influential variable is the drug price for combination therapy
- See the video “Tornado diagram” to learn how to do this graph in Excel

Tornado diagram in Excel

- A tornado diagram is a **essentially a bar chart**
- Select the upper and lower bound ICER estimates, click insert and select a 2-D horizontal bar chart
- Select the horizontal axis (where the ICER values are) and find the option "Vertical axis crosses." Choose Axis value and enter the baseline ICER 11499
- Select one bar and find the option that says Series Overlap and increase the overlap to 100
- That's it.
- You can then change the colors, labels, etc

Why are they useful? Department of Big Pictures

- Let's go to basics. You (somehow) come up with a range of plausible values. Then you calculate ICERs for the range. Then you graphically see which variable has the largest range (the “top” of the tornado)
- This way you identify –and communicate– the most important (**sensitive to changes**) variables in your model, in the sense that changing the value of those variables changes the ICER (or any other outcome of interest, but mostly the ICER)
- About the “somehow” part. If, for example, the estimate comes from a study, you could use the 95% confidence interval
- If it's the price of a medication, you could use different prices: wholesale, what the VA pays, what an average hospital pays
- You could ask a bunch of experts... **There are no hard rules about how one come up with these numbers**

The end of deterministic sensitivity analysis

- That's all about deterministic sensitivity analysis. Same rules apply to changing two variables at the same time, although it's harder to communicate what happens
- I covered **two-way sensitivity** analyses in the decision tree class
- Sometimes you want to do a **scenario analysis**
- It could involve changing several parameters at the same time. Say, if the cost of the combo medication is reduced by 25% and the government manages to reduce hospital costs, would the new drug be cost effective?
- This is a preview of our policy class, but one reason drug prices (technology in general) are cheaper in countries that use cost-effectiveness is because companies could do a CEA study and figure out at which price the new drug or technology will be cost effective, which means that it's more likely to be approved. In the US, pricing is more or less like this: *how high can we get away with?*

Probabilistic Sensitivity Analysis (PSA)

- Instead of changing one or more variables at a time we will change *all (or many) variables simultaneously* using a simulation (sometimes called **second-order Markov simulation**)
- Instead of coming up with a range of possible values for each variable we will assume that a parameter has a **probability distribution**
- Remember that a probability distribution assigns a probability to each possible outcome of a random variable
- We can **describe** a distribution using its parameters
- For example, a normal distribution can be described by its **mean** and **variance** (two-parameters)
- Another way: assuming a probability distribution for each parameter is equivalent to assuming many, many possible values a parameter could take (as many as infinity in most cases)

Stats 101 to understand the logic of PSA

- **Confidence intervals:** If we could repeat the experiment many, many times, 95 percent of the time the value of the parameter would be contained within the upper and lower bound of the confidence interval
- In the **frequentist** framework, the confidence interval is not a “probability” interval in the sense that the confidence interval is NOT defined as saying that we are “95 percent confident” that the true value of the parameter will be within the confidence interval
- That’s a bit closer to the Bayesian interpretation of confidence intervals (aka credible intervals)
- Language is so important. How we call things changes how we understand them. And it can lead to confusion

Stats 101 to understand the logic of PSA

- Currently the fights between Bayesians and Frequentists have mostly turned into love affairs
- Under flat priors, both produce similar results so no need to argue so much about which approach to statistics is better
- Flat priors mean that we approach an analysis assuming that we don't know anything about the true value of a parameter
- But the idea that we repeat the experiment many times to come up with a confidence interval is key to make sense of probabilistic sensitivity analysis
- **So what is PSA?** A way to come up with a confidence interval for a CEA study by “repeating” the study many many times (with a simulation) to understand the possible values the ICER could take – and thus quantify “uncertainty” (of the parameters)

Digression: Hypothesis testing

- There is a **direct connection between confidence intervals and hypothesis testing** you learned in stats 101
- Suppose we know that the 95% confidence interval for the relative risk is (0.4 - 0.6) and we want to test the hypothesis that the relative risk is zero. So the null is $H_0 = 0$. Would we reject the null?
- Yes, of course. Because if we repeat the experiment many times, 95% percent of the time zero is **not within the confidence interval** (thus the error is 5%)
- What about testing if the relative risk is 0.5? We would not reject the null
- **“Counting”** and repeating experiments is fundamental to statistics. Think about calculating the probability of obtaining a 2 rolling a dice. We calculate that by rolling the dice many times and counting the times 2 comes up

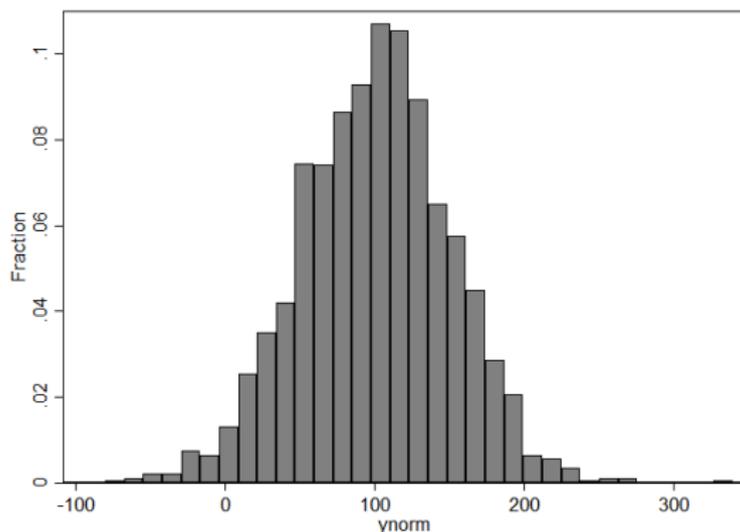
PSA and Groundhog Day

- Think of PSA as a kind of Groundhog Day experiment
- Bill Murray's character, Phil, keeps repeating the same day. Many times. Let's say 1000 times (we are never told how many times)
- He is trying to get Andie MacDowell's character, Rita, but he gets rejected and slapped pretty much all the time. Being a Hollywood movie, he does succeed – once (after he becomes a nice person, of course)
- We could use this “experiment” to quantify the probability that Phil gets Rita or the uncertainty: 1 in 1000 times
- Everybody's behavior in the movie kept changing with every new day, within certain parameters. But only once he succeeded

Simulation or how we get plausible numbers

- If we assume that a variable (in this case, a model parameter) has a distribution and we have the parameters of the distribution, we can use software to generate a sample of possible values
- Think of this as asking the computer to pick random numbers for you. The only condition is that these numbers must come from a probability distribution with certain characteristics
- For example, I used Stata to simulate 10 draws of a normal distribution with mean of 100 and standard deviation of 50. The first ten numbers were:
52.3, 125.6, 187.7, 113.5, 113.8, 138.1, 74.3, 77.4, -21.8, 37.0
- Pay attention to the numbers. All are reasonably close to 100, the mean. How close depends on the standard deviation, in this case, 50
- **Digression:** technically, computers generate **pseudo-random** numbers; need a “seed”

Normal distribution



- Simulated values for a normal variable with mean 100 and standard deviation of 50; $N(100, 50^2)$

PSA

- The procedure to conduct PSA is fairly simple:
 - 1 Assign a probability distribution to each model parameter
 - 2 Ask the computer to draw a number from these distributions
 - 3 Calculate ICER (or any other outcome of interest but mostly the ICER)
 - 4 Save the new ICER
 - 5 Repeat procedure many times (1000 or 10,000 times)
 - 6 Plot all ICERs in a cost-effectiveness plane
 - 7 Calculate statistics (like the percent of time the ICER is below threshold)
- The resulting “cloud” of ICERs will give you **a sense of the overall uncertainty of the study**

PSA for HIV example

- Campbell et al (2014) performed a PSA of the HIV example
- Assumed distributions

Table 1 Input Parameters and Ranges (Preliminary Model)

	Mean	Standard Error	Distribution	95% Interval for Inputs		Source
				Low	High	
Costs (1994–1995 British pounds)						
Direct medical costs State A	£1701	£1701	$\gamma(1, 1701)$	£42	£6360	a
Direct medical costs State B	£1774	£1774	$\gamma(1, 1774)$	£47	£6541	a
Direct medical costs State C	£6948	£6948	$\gamma(1, 6948)$	£175	£24,776	a
Community care costs State A	£1055	£1055	$\gamma(1, 1055)$	£23	£3894	a
Community care costs State B	£1278	£1278	$\gamma(1, 1278)$	£31	£4728	a
Community care costs State C	£2059	£2059	$\gamma(1, 2059)$	£49	£7577	a
Utilities						
Utility score State A	0.880	0.132	$\beta(5.33, 0.73)$	0.553	0.999	b
Utility score State B	0.780	0.117	$\beta(9.78, 2.76)$	0.527	0.954	b
Utility score State C	0.700	0.105	$\beta(13.33, 5.71)$	0.486	0.877	b
Combination therapy treatment effect						
Relative risk ^c	0.509	0.088	Log-normal(-0.68, 0.17)	0.365	0.711	a

a. Chancellor and colleagues⁸ and Briggs and colleagues.⁵

b. Kauf and colleagues¹⁰ and Tsevat and colleagues.¹¹

c. Assumed to only apply to the first 2 years of the model (consistent with Chancellor and colleagues⁸).

PSA for HIV example

- Performed 10,000 simulation draws

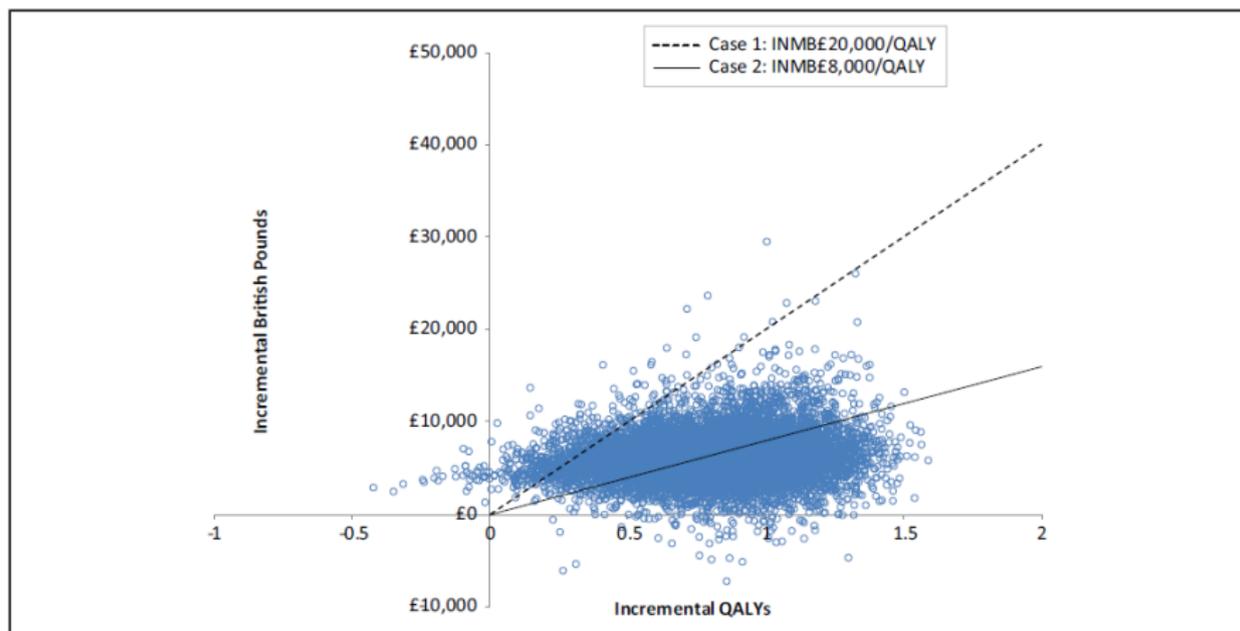


Figure 1 Cost-effectiveness plane with 10,000 Monte Carlo draws and willingness-to-pay lines for preliminary model Case 1 ($\lambda = £20,000/QALY$) and Case 2 ($\lambda = £8,000/QALY$).

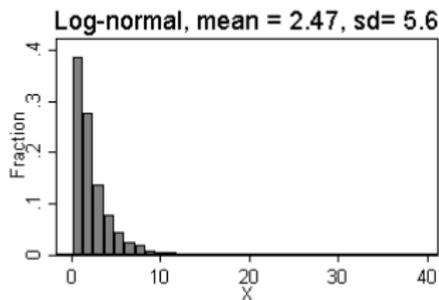
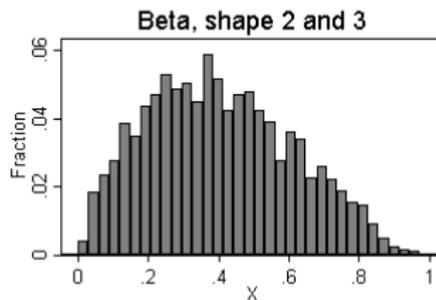
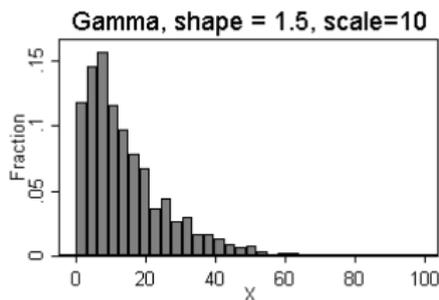
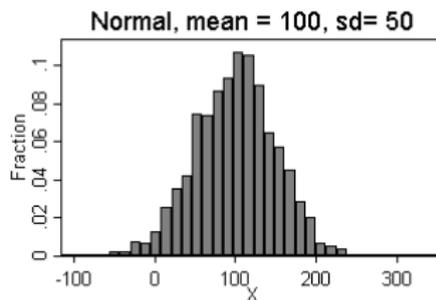
Why is PSA important?

- It gives you an idea of the **overall uncertainty** of the study
- If the points are scattered all over the four quadrants, there is a lot of uncertainty
- In general results are presented as the percent of times the ICER is below the threshold line
- Obviously, less uncertainty means that our model is less sensitive to the value of parameters, which is good
- PSA is essentially trying to mimic frequentist **statistical inference**

Choosing probability distributions for probabilistic sensitivity analysis

- The normal distribution is the first one you learn but in CEA few parameters distribute normal
- Costs, for example, cannot distribute normal because they have to be > 0 and costs are usually skewed (but prices are often close to normally distributed)
- Utilities and probabilities must be bounded between 0 and 1 so they can't be normally distributed
- Common distributions in PSA are Gamma, Beta, and Log-normal, Poisson depending on the parameter
- Obviously, the domain or support of the distribution must match the characteristics of the parameter. It would be a bad idea to use a normal distribution to model the number of test needed, which are integer, positive numbers like 0, 2, 3, etc

Example distributions



■ See Stata code for simulations

Characteristics

- **Gamma:** $\gamma(k, \theta)$ has two parameters (shape and scale), support from > 0 to infinity. Used to simulate costs because costs are positive and skewed
- **Beta:** $\beta(a, b)$ has two shape parameters, support from 0 to 1. Good for simulating utilities or probabilities because they are bounded between 0 and 1
- **Log-normal:** the exponent of a normal distribution $\exp(N(\mu, \sigma^2))$. Support from 0 to infinity
- Note: support is also called the **domain** of a distribution; the range of values for which the distribution is defined

How to simulate numbers

- Most statistical packages (SAS, Stata, R, Excel) have random numbers generators for common distributions
- **Inverse Transform Method:** It is possible to simulate numbers from any (continuous) distribution by using a uniform random generator
- We only need 1) a random generator for the uniform distribution and 2) the formula of the inverse cumulative distribution (may not exist in closed form)
- Now most software has random number generators so no need for the inverse transform method, but you still see this because that's how some people learned how to simulate random numbers
- See Excel file for examples

Summary

- We have used Markov models as examples, but everything we have learned apply to decision trees
- Sensitivity analyses are vital parts of CEA
- PSA may be complicated to implement but it is very easy to understand