# UGA L3 MIASH: Econometrics 1
# Review of Probability

Michal W. Urdanivia*

*Université de Grenoble Alpes, Faculté d'Économie, GAEL,
e-mail: michal.wong-urdanivia@univ-grenoble-alpes.fr

September 26, 2017

## References

- Wooldridge (2013) appendix B
- Stock and Watson (2009) chapter 2
- Abbring (2001) sections 2.1-2.3
- Diez, Barr, and Cetinkaya-Rundel (2012) chapter 2
- Grinstead and Snell (2003) chapters 1-7

- These lectures are mostly based on Wooldridge, so that will generally be the best reference to read.
- The others are some freely available alternatives that cover similar material.
- Abbring (2001) is most similar to Wooldridge in terms of length and its focus on econometrics.
- Diez, Barr, and Cetinkaya-Rundel (2012) and Grinstead and Snell (2003) also cover similar material, but explain things in more depth.

# Probability

- Purpose: system for quantifying chance and making predictions about future events
- Interpretations:
  - Relative frequency of an event in many repeated trials
  - Subjective assessment of the likelihood of an event

UNIVERSITÉ Grenoble Alpes

## Basic definitions

- The properties of probability are fairly intuitive, and you can work with probabilities without worrying too much about how probability is formally defined.
- However, if we want to be mathematically rigorous, we should carefully define what we mean by probability.
- This slide formally defines probability. Do not worry if it seems very abstract.

# Basic definitions

- Random experiment: procedure that has well-defined set of outcomes and "could" be infinitely repeated
- Sample space: set of possible outcomes of an experiment, $S = \{a_1, a_2, ..., a_J\}$
- Event: any subset of sample space, $A \subseteq S$
- Probability: function from all subsets of the sample space, $S$, to $[0, 1]$ such that
    1. $\mathbb{P}(S) = 1$
    2. $1 \geq \mathbb{P}(A) \geq 0$ for all $A \subseteq S$
    3. If $A_1, A_2, ...$ are disjoint events then $\mathbb{P}(A_1 \cup A_2 \cup ...) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + ...$
- Random variable: function from $S$ to a number

# Basic definitions

- An example of a random experiment is rolling a dice. In this case the sample space is $S = \{1, 2, 3, 4, 5, 6\}$.
- Some events are: get a $1 = \{1\}$, get more than 4, $\{5, 6\}$, etc.
- A random variable is a numeric representation of a random experiment.
- For example, our experiment could be flipping a coin. Then $S = \{\text{heads}, \text{tails}\}$.
- A function assigns a number to each element of $S$, so one example of a random variable is $X = \begin{cases} 1 & \text{if heads} \\ 0 & \text{if tails} \end{cases}$.
- Another is $Y = \begin{cases} 50 & \text{if heads} \\ -257 & \text{if tails} \end{cases}$.
- A third example is $Z = \begin{cases} 1 & \text{if heads} \\ 1 & \text{if tails} \end{cases}$.

# Properties of probability

1. $\mathbb{P}(\emptyset) = 0$
2. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
3. $A \subseteq B$ implies $\mathbb{P}(A) \leq \mathbb{P}(B)$
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

# Conditional probability

- Conditional probability: $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$
  - Satisfies axioms of unconditional probability (and so also has properties on previous slide)
- Bayes' Rule $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$
- $A$ is independent of $B$ iff $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, denote as $A \perp\!\!\!\perp B$
  - $A \perp\!\!\!\perp B$ implies $\mathbb{P}(A|B) = \mathbb{P}(A)$ and $\mathbb{P}(B|A) = \mathbb{P}(B)$.

## Discrete random variables

- Recall: a random variable is a function from $S$ to a number
- A random variable is discrete if it can only take countably many different values
- E.g. $X \in \{x_1, ..., x_m\}$
- Probability mass function (PMF) of $X$, $p_i = \mathbb{P}(X = x_i)$ gives the probability that $X$ equals each of its possible values
- Cumulative distribution function (CDF):
  $F(x) = \mathbb{P}(X \leq x) = \sum_{i=1}^{m} p_i 1\{x_i \leq x\}$

## Continuous random variables

- A random variable is continuous if it takes on any single real value with zero probability
- Set of possible values is uncountably infinite (e.g. real line or line segment)
- CDF is continuous and differentiable
- Probability density function (PDF): the derivative of the CDF

$$f(x) = \frac{dF}{dx}(x) \text{ and } F(x) = \int_{-\infty}^{x} f(t)dt$$

# Bivariate distributions

- Discrete:
  - Joint PMF $f_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y)$
  - Marginal PMF $f_X(x) = \mathbb{P}(X = x) = \sum_{y \in \text{all values of } Y} \mathbb{P}(X = x, Y = y)$
  - Conditional PMF $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
- Continuous:
  - joint CDF $F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,t)dudt$
  - joint PDF $f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y)$
  - Marginal CDF for $X$, $F_X(x) = \mathbb{P}(X \leq x) = F_{X,Y}(x, \infty)$
  - Marginal PDF for $X$, $f_X(x) = \frac{dF_X}{dx}(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$
  - Conditional PDF $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

# Independence

- Random variables $X, Y$ are independent if $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all $x, y$
- $X \perp\!\!\!\perp Y$ implies $f_{X|Y}(x|y) = f_X(x)$

## Expectation

- The expectation of $X$ is

$$\mathbb{E}[X] = \begin{cases} \sum_{i=1}^{m} x_i f_X(x_i) & \text{if } X \text{ discrete} \\ \int_{-\infty}^{\infty} x f_X(x) dx & \text{if } X \text{ continuous} \end{cases}$$

  - Is a constant
  - a.k.a. expected value, population average, population mean, first moment
- Expectation is linear:

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

- Expectation of function $g$:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

# Variance and other moments

- $k$th moment of $X$: $\mathbb{E}[X^k]$
- $k$th central moment of $X$: $\mathbb{E}\left[(X - \mathbb{E}[X])^k\right]$
- Variance is the 2nd central moment

$$\begin{aligned}\mathbb{V}(X) =& \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\ =& \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

- $\mathbb{V}(a + bX) = b^2 \mathbb{V}(X)$
- Standard deviation is $\sqrt{\mathbb{V}(X)}$

# Covariance

- Covariance of $X$ and $Y$:

$$\begin{aligned}\mathbb{C}(X,Y) =& \mathbb{E}\left[(X-\mathbb{E}[X])\,(Y-\mathbb{E}[Y])\right]\\ =& \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]\\ =& \mathbb{E}[(X-\mathbb{E}[X])\,Y] = \mathbb{E}[X\,(Y-\mathbb{E}[Y])]\end{aligned}$$

- $\mathbb{C}(X,X) = \mathbb{V}(X)$
- $\mathbb{C}(a_1 + b_1X + c_1Y, a_2 + b_2X + c_2Y) = b_1b_2\mathbb{V}(X) + c_1c_2\mathbb{V}(Y) + (b_1c_2 + c_1b_2)\mathbb{C}(X,Y)$
- $\mathbb{V}(a + bX + cY) = b^2\mathbb{V}(X) + c^2\mathbb{V}(Y) + 2bc\mathbb{C}(X,Y)$
- $\mathbb{V}\left(\sum_{i=1}^N b_iX_i\right) = \sum_{i=1}^N \left(\sum_{j=1}^N b_ib_j\mathbb{C}(X_i,X_j)\right)$

## Correlation

- Correlation of $X$ and $Y$ is $\mathrm{Corr}(X, Y) = \frac{\mathbb{C}(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}$
- Cauchy-Schwartz inequality: $|\mathbb{C}(X, Y)| \leq \sqrt{\mathbb{V}(X)\mathbb{V}(Y)}$
  - Implies $-1 \leq \mathrm{Corr}(X, Y) \leq 1$
- $\mathrm{Corr}(X, Y) = \pm 1$ iff $Y = a + bX$

# Conditional expectation

- Conditional expectation of $Y$ given $X = x$:

$$\mathbb{E}[Y|X=x] = \begin{cases} \sum_{i=1}^{m} y_i f_{Y|X}(y_i|x) & \text{if discrete} \\ \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy & \text{if continuous} \end{cases}$$

- $\mathbb{E}[Y|X]$ is a function of $X$
- Has all properties of unconditional expectation
- Properties:
  - $\mathbb{E}[g_1(X) + g_2(X)Y|X] = g_1(X) + g_2(X)\mathbb{E}[Y|X]$
  - Law of iterated expectations: $\mathbb{E}_X\left[\mathbb{E}_{Y|X}[Y|X]\right] = \mathbb{E}[Y]$
  - $\mathbb{E}\left[X\left(Y - \mathbb{E}[Y|X]\right)\right] = 0$

# Conditional variance

- Conditional variance: $\mathbb{V}(Y|X) = \mathbb{E}\left[(Y - \mathbb{E}[Y|X])^2 \, |X\right]$
- Relation to variance:

$$\mathbb{V}(Y) = \mathbb{E}_X\left[\mathbb{V}(Y|X)\right] + \mathbb{V}\left(\mathbb{E}[Y|X]\right)$$

  - Analysis of variance (ANOVA)
  - $\mathbb{E}_X\left[\mathbb{V}(Y|X)\right]$ is within-$X$ variance
  - $\mathbb{V}\left(\mathbb{E}[Y|X]\right)$ is between-$X$ variance

# References

Abbring, Jaap. 2001. "An Introduction to Econometrics: Lecture notes." URL http://jabbring.home.xs4all.nl/courses/b44old/lect210.pdf.

Diez, David M, Christopher D Barr, and Mine Cetinkaya-Rundel. 2012. *OpenIntro Statistics*. OpenIntro. URL http://www.openintro.org/stat/textbook.php.

Grinstead, Charles M and J. Laurie Snell. 2003. *Introduction to Probability*. American Mathematical Society. URL http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/book.html.

Stock, J.H. and M.W. Watson. 2009. *Introduction to Econometrics, 2/E*. Addison-Wesley.

Wooldridge, J.M. 2013. *Introductory econometrics: A modern approach*. South-Western.