# Probability

Paul Schrimpf

UBC
Economics 326

January 23, 2018

# References

- Wooldridge (2013) appendix B
- Stock and Watson (2009) chapter 2
- Linton (2017) chapters 1-5
- Abbring (2001) sections 2.1-2.3
- Diez, Barr, and Cetinkaya-Rundel (2012) chapter 2
- Grinstead and Snell (2003) chapters 1-7

# 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

# Section 1

# Definitions

# 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. $P(S) = 1$
  2. $1 \geq P(A) \geq 0$ for all $A \subseteq S$
  3. If $A_1, A_2, ...$ are disjoint events then $P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$
- Random variable: function from $S$ to a number

Section 2

# Properties

# Properties of probability

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

# Conditional probability

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

# Section 3

# Random variables

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

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

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

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

- Expectation of function $g$:

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

# Variance and other moments

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

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

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

# Covariance

- Covariance of $X$ and $Y$:

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

- $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(a_1 + b_1 X + c_1 Y, a_2 + b_2 X + c_2 Y) = b_1 b_2 \mathrm{Var}(X) + c_1 c_2 \mathrm{Var}(Y) + (b_1 c_2 + c_1 b_2)\mathrm{Cov}(X,Y)$
- $\mathrm{Var}(a + bX + cY) = b^2\mathrm{Var}(X) + c^2\mathrm{Var}(Y) + 2bc\mathrm{Cov}(X,Y)$
- $\mathrm{Var}\left(\sum_{i=1}^N b_i X_i\right) = \sum_{i=1}^N \left(\sum_{j=1}^N b_i b_j \mathrm{Cov}(X_i, X_j)\right)$

# Correlation

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

# Conditional expectation

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

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

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

# Conditional variance

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

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

  - Analysis of variance (ANOVA)
  - $\mathrm{E}_X\left[\mathrm{Var}(Y|X)\right]$ is within-$X$ variance
  - $\mathrm{Var}\left(\mathrm{E}[Y|X]\right)$ 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.

Linton, Oliver B. 2017. *Probability, Statistics and Econometrics*. Academic Press. URL http://gw2jh3xr2c.search.serialssolutions.com/?sid=sersol&SS_jc=TC0001868500&title=Probability%2C%20statistics%20and%20econometrics.

# References

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.