Quantitative Economics Lecture 2 - Probability Distributions

Richard Povey

University of Oxford

24th April 2018

richard.povey@hertford.ox.ac.uk, richard.povey@st-hildas.ox.ac.uk

With thanks to Ian Crawford, Margaret Stevens, Vanessa Berenguer-Rico.

Distribution Functions

Functions associated with Random Variables:

- Cumulative Distribution Function: $F_X(x)$
- Probability Density Function (continuous) or Probability Mass Function (discrete) : $f_X(x)$

Important - Note on Notation

- Upper case X for *random variables*.
- Lower case x or x_i for realisations / specific values.
- Subscripts / superscripts denote individual *observations* X_i.

Distribution Functions

Definition

The cumulative distribution function (CDF) of a random variable X, denoted by $F_X(x)$ is the function

$$F_X(x) = P_X(X \leq x) : \; -\infty < x < \infty$$

Properties:

- (i) $0 \leq F(x) \leq 1$ for all x
- (ii) $F(x)$ is non-decreasing in x
- (iii) $\lim_{x \to -\infty}\{F(x)\} = 0$ and $\lim_{x \to \infty}\{F(x)\} = 1$

Definition - Discrete Distribution PMF

The real-valued function $f_X(x)$ defined by $f_X(x) = P(X = x)$ is called the probability mass function (PMF) of X.

Distribution of a Discrete Random Variable

Example: Rolling two dice

Define X as the sum of the outcomes from each die:
$f_X(x) = P(X = x)$ for all x.

$X = x$	A	$f_X(x)$
2	$\{(1,1)\}$	$1/36$
3	$\{(1,2),(2,1)\}$	$2/36$
4	$\{(1,3),(2,2),(3,1)\}$	$3/36$
5	$\{(1,4),(2,3),(3,2),(1,4)\}$	$4/36$
6	$\{(1,5),(2,4),(3,3),(4,2),(5,1)\}$	$5/36$
7	$\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$	$6/36$
8	$\{(2,6),(3,5),(4,4),(5,3),(6,2)\}$	$5/36$
9	$\{(3,6),(4,5),(5,4),(6,3)\}$	$4/36$
10	$\{(4,6),(5,5),(6,4)\}$	$3/36$
11	$\{(5,6),(6,5)\}$	$2/36$
12	$\{(6,6)\}$	$1/36$

Distribution of a Continuous Random Variable

Definition

A continuous random variable X on a probability space $(X, \mathcal{A}, P)$ is a real-valued function $X(\omega)$, $\omega \in \Omega$, such that for $-\infty < x < \infty$, $\{\omega | X(\omega) \leq x\}$ is an event.

Definition

A probability density function (PDF) is a non-negative function f such that $\int_{-\infty}^{\infty} f_X(y)dy = 1$.

- Cumulative Distribution Function:

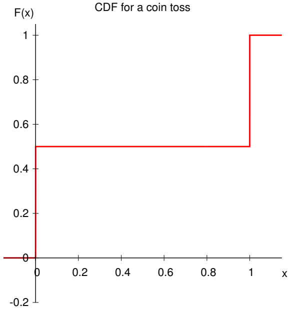
$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy \ : \ -\infty < x < \infty.$$

  - **Remark 1**: It is possible to construct continuous distribution functions that do not have densities.
  - **Remark 2**: Distribution functions that have densities are called absolutely continuous (the usual case).

Cumulative Distribution Functions

Example: Discrete Random Variable. Tossing a coin

$$F_X(x) = \begin{cases} 0 & -\infty < x < 0 \\ 1/2 & 0 \leq x < 1 \\ 1 & 1 \leq x < \infty \end{cases}$$



Cumulative Distribution Functions

Example: Choose a number from the interval $[a, b]$ at random

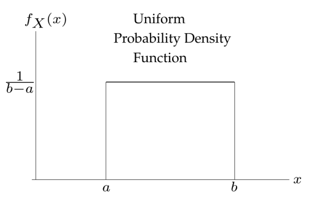
Continuous Random Variable. Uniform CDF:

$$F_X(x) = \left\{ \begin{array}{cc} 0 & -\infty < x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & b \leq x < \infty \end{array} \right\}$$

Uniform PDF:

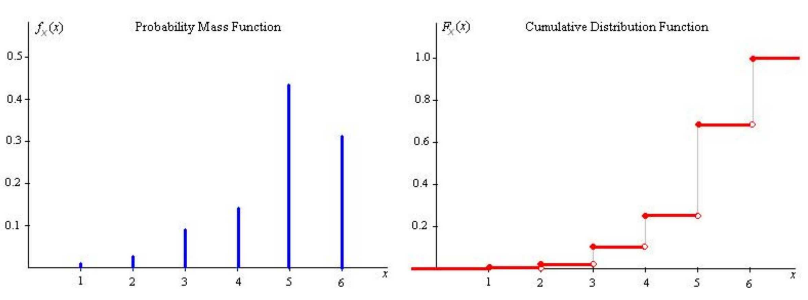
$$f_X(x) = \frac{1}{b-a} \text{ for } a < x < b$$

and 0 elsewhere



Discrete Random Variables

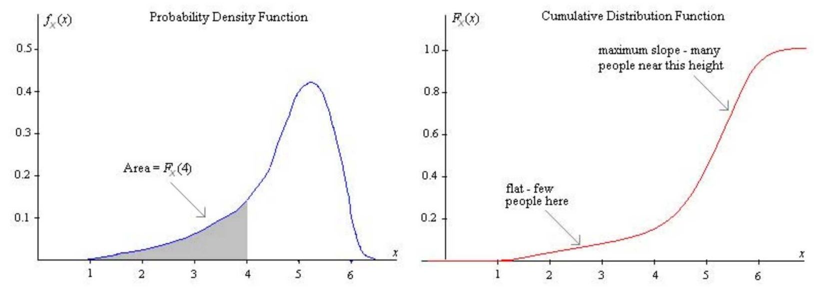
Example: X=Height of person to nearest foot



$$\sum_{x=1}^{6}[f_X(x)] = 1 \qquad F_X(x) = \sum_{y\leq x}[f_X(y)]$$

Continuous Random Variables

Example: X=Exact height of person



$f_X(x) = \frac{dF_X(x)}{dx}$ $\quad \int_{-\infty}^{\infty} f_X(x)dx = 1$ $\quad F_X(x) = \int_{-\infty}^{x} f_X(y)dy$

Bernouilli Distribution

Special Distributions

- **Discrete r.v.** Main Distributions used for QE course: Bernoulli, Binomial (Other examples: Poisson)
- **Continuous r.v.** Main Distributions for QE: Uniform, Normal, Chi-Square, F, Student's t (Other examples: Logistic)

Bernoulli Distribution : $X \overset{d}{=} B(p)$ (or $X \sim B(p)$)

- Binary (dichotomous) outcomes: 0 (failure) and 1 (success). Example: Is gender of a new born female? (X=1 yes, X=0 no)

$$X = \left\{ \begin{array}{cc} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{array} \right\}$$

- PMF: $f_X(x) = p^x(1-p)^{1-x}$ for $x \in \{0, 1\}$ and 0 elsewhere.

Binomial Distribution

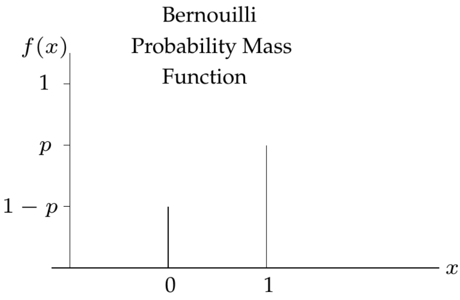
Binomial Distribution : $Y \stackrel{d}{=} B(n,p)$ (or $Y \sim B(n,p)$)

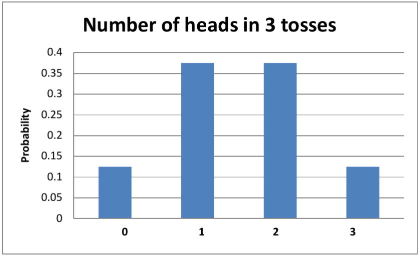
- Total number of successes from n repetitions of the same Bernoulli experiment. So $Y = \sum_{i=1}^{n} [X_i]$ with each $X_i \stackrel{d}{=} B(p)$. (Important for Lecture 3.)
- Binomial random variable $X \stackrel{d}{=} B(n,p)$ takes values $\{0, 1, 2, ..., n\}$
- PMF: $f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x \in \{0, 1, 2, ..., n\}$ and 0 elsewhere.
- Look in statistics textbooks for Binomial probabilities. These can get complicated. Fortunately, as we see in Lecture 2, mostly in QE we rely on n being high and a *normal approximation* to the $B(n,p)$ distribution.

Binomial and Bernouilli PMF Examples

$X \stackrel{d}{=} B(p)$ and $Y \stackrel{d}{=} B(3, 0.5)$

Y=y	$f_Y(y)$
0	0.125
1	0.375
2	0.375
3	0.125





$$P(\{(HHH)\}) = P(\{(TTT)\}) = \left(\tfrac{1}{2}\right)^3 = 0.125$$
$$P(\{(HHT),(HTH),(THH)\}) = P(\{(TTH),(THT),(HTT)\}) = 3\left(\tfrac{1}{2}\right)^3 = 0.375$$

Characterising a Distribution

- Centrality: **Expected Value**
- Dispersion: **Variance, Standard Deviation**
- Asymmetry: **Skewness**
- Tailedness: **Kurtosis**

Expected Value

Definition

- Let X be a discrete random variable having mass function $f_X(x)$. If $\sum_{j=1}^{\infty} [|x_j| f_X(x_j)] < \infty$, we say that X has finite expectation and we define its expectation by:

  $E[X] = \sum_{j=1}^{\infty} [x_j f_X(x_j)] = \sum_{j=1}^{\infty} [x_j P(X = x_j)].$

  If $\sum_{j=1}^{\infty} [|x_j| f_X(x_j)] = \infty$ then we say that X has no finite expectation and $E[X]$ is undefined.

Definition

- Let X be a continuous random variable having density $f_X(x)$. If $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$, we say that X has finite expectation and we define its expectation by:

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

Characterising a Distribution: Expected Value

Example: Bernoulli distribution

$$X \overset{d}{=} B(p) \Longrightarrow E[X] = \sum_{i=1}^{2} [x_i P(X = x_i)]$$

$$= \ 1 \times P(X = 1) + 0 \times P(X = 0) = P(X = 1) = \boxed{\text{p}}$$

Example: Uniform distribution

$$X \overset{d}{=} U[a, b] \Longrightarrow E[X] = \int_a^b x \left(\frac{1}{b-a}\right) dx$$

$$= \left[\left(\frac{1}{b-a}\right)\frac{x^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{(a+b)(b-a)}{2(b-a)} = \boxed{\frac{a+b}{2}}$$

Characterising a Distribution: Expected Value

Properties of the Expected Value

- If c is a constant and $P(X = c) = 1$ then $E[X] = c$.
- If c is a constant and $P(X = c) = 1$ then $E[g(X)] = g(c)$.
- If b and c are constants then $E[b + cX] = b + cE[X]$. (*Linearity* of expectation operator.)
- **Jensen's Inequality**:
  - If X is a random variable and $g(X)$ is a concave function, then $E\left[g(X)\right] \leq g\left(E[X]\right)$.
  - If $g(X)$ is a convex function, then $g\left(E[X]\right) \leq E\left[g(X)\right]$.
  - For *strict* concavity / convexity (and non-constant X), $\leq$ becomes $<$ in each of the above inequalities.

Jensen's Inequality

Example: Core Micro: Expected utility theory

- If X is a randomly distributed amount of wealth (i.e. a lottery) and utility $u(X)$ is a *strictly concave* function (due to diminishing marginal utility) then:

$$u(E(X)) > E[u(X)]$$

- So, utility from expected wealth is greater than expected utility of wealth.

- The agent with concave utility would be better off with the expected value of the lottery for certain than with the risky lottery (i.e. they are risk averse).

Example: Core Macro: Precautionary saving

- Suppose a consumer has logarithmic utility $u(C) = \ln(C)$ so MU is $u'(C) = \frac{1}{C}$.
- If the consumer's subjective discount rate equals the interest rate then the Euler equation for optimal consumption is $u'(C_1) = E[u'(C_2)]$.
- $u'(C)$ is *strictly convex*, so under uncertainty about C $E[u'(C)] > u'(E[C])$. Hence the Euler equation implies that:

$$u'(C_1) > u'(E[C_2]) \Longrightarrow \frac{1}{C_1} > \frac{1}{E[C_2]} \Longrightarrow E[C_2] > C_1$$

- The consumer expects/plans to consume *more tomorrow* (whereas with *certainty* about the future we get $C_2 = C_1$).

Characterising a Distribution: Variance

Definition

- The **variance** is simply an expectation - the expectation of the squared difference between the random variable and its mean/expected value:

$$\sigma_X^2 = Var(X) = E\left[(X - E[X])^2\right]$$

- Discrete Random Variables:

$$\sigma_X^2 = Var(X) = \sum_{i=1}^{k}\left[(x_i - E[X])^2 f_X(x_i)\right]$$

- Continuous Random Variables:

$$\sigma_X^2 = Var(X) = \int_{-\infty}^{\infty}(x - E[X])^2 f_X(x)dx$$

Characterising a Distribution: Variance

Properties of the Variance

- $Var(X) = E[X^2] - (E[X])^2$
- $Var(cX) = c^2 Var(X)$
- $Var(b + cX) = c^2 Var(X)$
- By Jensen's inequality, for a non-constant X, $Var(X) > 0$ since $(X - E[X])^2$ is a strictly convex function.

Definition

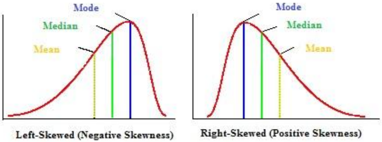
Standard Deviation: $\sigma_X = sd(X) = \sqrt{Var(X)}$

Characterising a Distribution: Skewness and Kurtosis

Asymmetry: **Skewness**

$$Skew(X) = E\left[\left(\frac{X - E(X)}{\sigma_X}\right)^3\right] = \frac{E[(X - E[X])^3]}{Var(X)^{\frac{3}{2}}}$$

- Skewness is positive if there is a "long right tail" (e.g. income distribution) and negative if there is a long left tail (e.g. age of death).



Source: http://www.statisticshowto.com

Tailedness: **Kurtosis**

$$Kurt(X) = E\left[\left(\frac{X - E(X)}{\sigma_X}\right)^4\right] = \frac{E[(X - E[X])^4]}{Var(X)^2}$$

Characterising a Distribution: Bernoulli Distribution

Example: Bernouilli Distribution $X \stackrel{d}{=} B(p)$

- $E[X] = P(X=1) = \boxed{\text{p}}$

- $Var(X) = \sum_{i=1}^{2}\left[(x_i - E[X])^2 f_X(x_i)\right] =$
  $(1-p)^2 p + (0-p)^2(1-p) = p(1-p)((1-p)+p) = \boxed{p(1-p)}$

- $Skew(X) = \frac{(1-p)^3 p + (0-p)^3(1-p)}{(p(1-p))^{\frac{3}{2}}} = \frac{p(1-p)((1-p)^2 - p^2)}{(p(1-p))^{\frac{3}{2}}} =$
  $\frac{p(1-p)(1-2p)}{(p(1-p))^{\frac{3}{2}}} = \boxed{\frac{(1-2p)}{(p(1-p))^{\frac{1}{2}}}}$

- We can see that if $p < 0.5$ then the B(p) Bernoulli distribution is positively skewed and if $p > 0.5$ then it is negatively skewed.