1 Discrete random variables

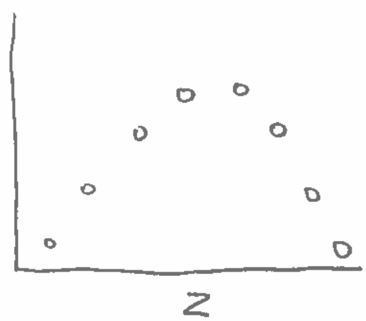
- What is a random variable? It is a quantity that has values determined by chance.

- What is a discrete random variable z? If we specify 2 values $a, b, a < b$ then there are a finite number of values of z between a and b. Not necessarily integers. For example, the random variable "the age of students in years ESS 575 to the nearest 1/2 year, the random variable is discrete but it has fractional values, i.e., 27.5. Most often we will use counts, which are strictly non-negative integers.

- Probability mass function, z, PMF (also called probability function, discrete destiny function)

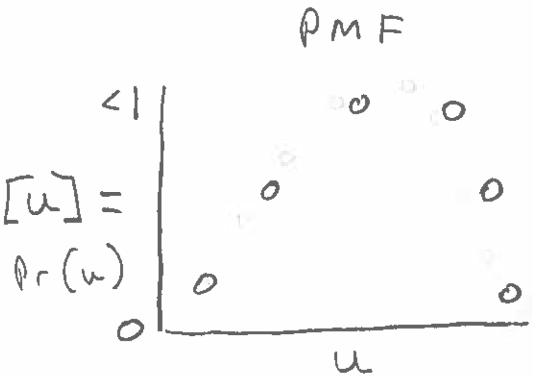
  - notation $[z]$— expand on notation of others, $f(Z = z)$
  - $[z]$ is a function that returns the probability of a specific value z of the random variable.
  - Support of random variable z is defined as all values of z for which $[z] > 0$ and defined.

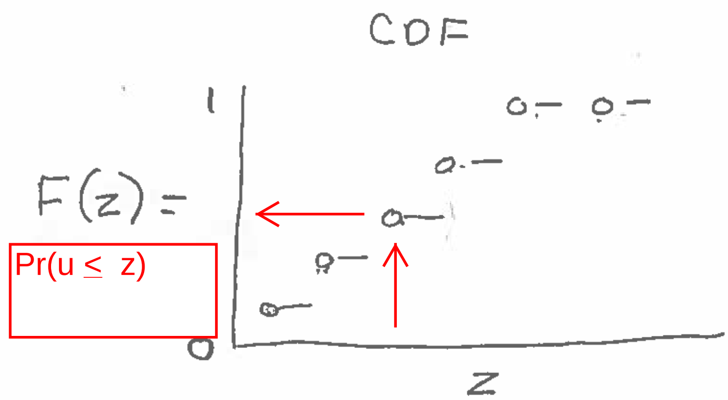
$[z] = Pr(z)$



  - requirements to be a PMF
    * $[z] \geq 0$
    * $\sum_{z \in s} z = 1$, where s is the support of the random variable
  - moments of PMF
    * first moment, the expected value (or mean) $= \text{E}(z) = \mu = \sum_{z \in s} z[z]$, approximated from many (n) random draws from $[z]$ using $\text{E}(z) \simeq \frac{1}{n} \sum_{i=1}^{n} z_i$
    * second central moment, the variance $= \text{E}(z - \mu)^2 = \sigma^2 = \sum_{z \in s}(z - \mu)^2[z]$, approximated from many (n) random draws from $[z]$ using $\text{E}(z - \mu)^2 \simeq \frac{1}{n} \sum_{i=1}^{n}(z_i - \mu)^2$

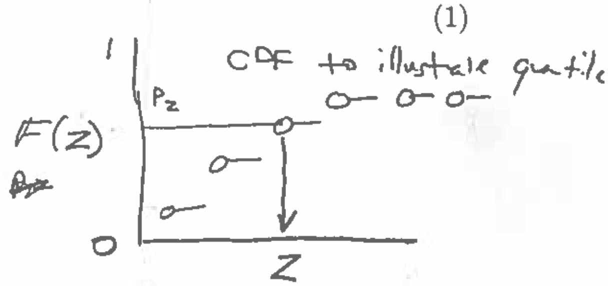
- cumulative distribution function for z
  - Plots here, defining $[u]$ as pmf



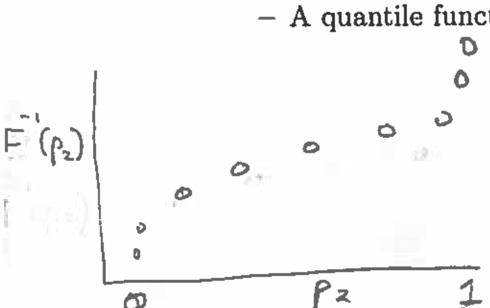


  - A cumulative distribution function for the random variable z is

$$F(z) = \sum_{u<z} [u] \tag{1}$$



- quantile function for z
  - A quantile function for the random variable z is



$$F(z) = p_z \quad = \quad \sum_{u<z} [u] \tag{2}$$

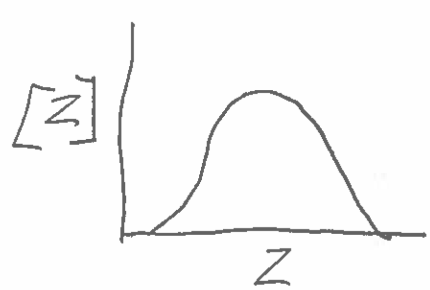
$$\cancel{F^{-1}(p_z) \quad = \quad F(z) \leq p_z} \tag{3}$$

$$F^{-1}(p_z) = z \text{ for which } F(z) \leq p_z$$

2 Continuous random variables

If we specify two values a and b, a < b, then there are an infinite number of values of z in the interval a to b.

- Probability density function, PDF, $[z]$
  - notation $[z]$, $f(z)$, $z \sim \text{normal}()$
  - $[z]$ gives the *probability density* of a specific value of the random variable $= z$.
  - Support of random variable z is defined as all values of z for which $[z] > 0$ and defined.
  - requirements
    - $[z] \geq 0$
    - $\int_{-\infty}^{\infty} [z] dz = 1$
    - $\Pr(a < z < b) = \int_a^b [z] dz$



– More about probability density? Plot here

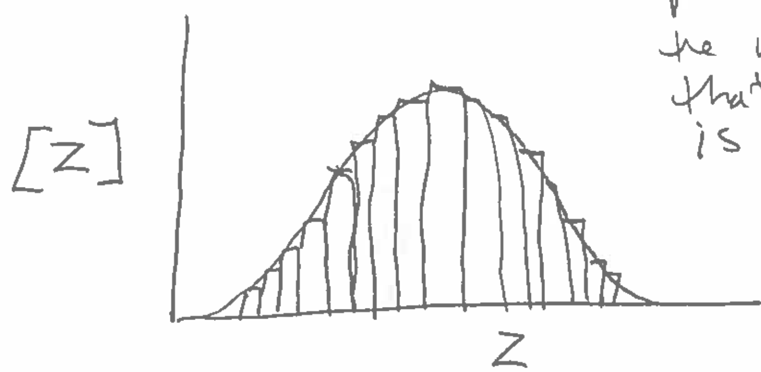
probality density is the value of z such that the third requirement is true:.

$$\int_{-\infty}^{\infty} [z]\,dz = 1$$

dist ↓

height of bars $= \left[\frac{a+b}{2}\right]$

width of bars $= \Delta z = b-a$

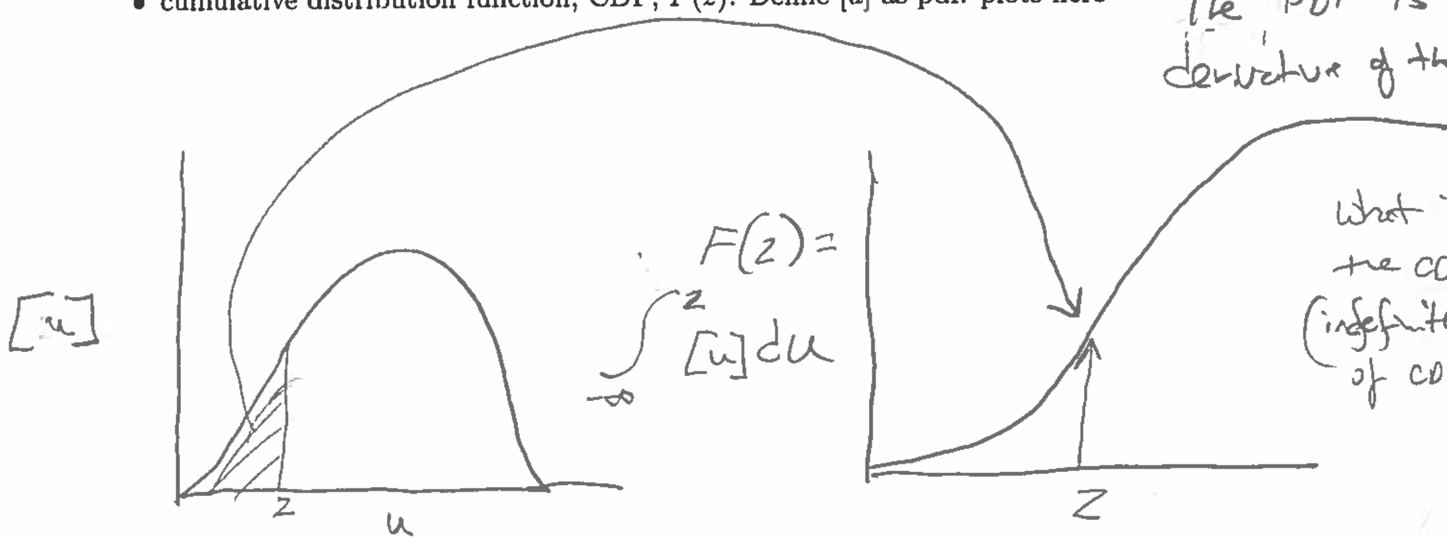


– moments

* first moment, the expected value (or mean) $= \mathrm{E}(z) = \mu = \int_{-\infty}^{\infty} z[z]dx$, approximated from many (n) random draws from $[z]$ using $\mathrm{E}(z) \simeq \frac{1}{n}\sum_{i=1}^{n} z_i$
* second central moment, the variance $= \mathrm{E}(z-\mu)^2 = \sigma^2 = \int_{-\infty}^{\infty}(z-\mu)^2[z]dx$, approximated from many (n) random draws from $[z]$ using $\mathrm{E}(z-\mu)^2 \simeq \frac{1}{n}\sum_{i=1}^{n}(z_i-\mu)^2$

• cumulative distribution function, CDF, $F(z)$. Define $[u]$ as pdf. plots here

The PDF is the derivative of the CDF



What is the CDF? (indefinite int of CDF)

Do curves without areas first.

$$F(z) = \int_{-\infty}^{z} [u]du \tag{4}$$

- quantile function

$$F(z) = p_z = \int_{-\infty}^{z} [u] du \tag{5}$$

$$F^{-1}(p_z) = F(z) \leq p_z \tag{6}$$