Discrete random variables

**Definition.** Let \( \Omega \) be a sample space. A **discrete random variable** is a function \( X : \Omega \to \mathbb{R} \) that takes on a finite number of values \( a_1, a_2, \ldots, a_n \) or an infinite number of values \( a_1, a_2, \ldots \).

\[
p(a_i) > 0 \text{ for } i = 1, 2, \ldots \\
p(a) = 0 \text{ if } a \notin \{a_1, a_2, \ldots\} \\
\sum_i p(a_i) = 1
\]

**Definition.** The **probability mass function** \( p \) of a discrete random variable \( X \) is the function \( p : \mathbb{R} \to [0, 1] \), defined by

\[
p(a) = P(X = a) \quad \text{for } -\infty < a < \infty.
\]

- Support finite or countable \( \{a_1, \ldots, a_n, \ldots\} \)
  - \( p(a_i) > 0 \) for \( i = 1, 2, \ldots \)
  - \( p(a) = 0 \) if \( a \notin \{a_1, a_2, \ldots\} \)
  - \( \sum_i p(a_i) = 1 \)

- What happens when the support is uncountable? E.g., \([0, 1]\) or \( \mathbb{R}^+ \) or \( \mathbb{R} \)
  - \( p(a_i) \) must be 0 because \( |\mathbb{R}| = 2^{\aleph_0} > \aleph_0 = |\mathbb{N}| \)
  - hence, \( \sum_i p(a_i) = 0 \)
Continuous random variables

- We cannot assign a “mass” to a real number, but we can assign it to an interval!

\[ F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x) \, dx \] [Cumulative Distribution Function]
Density function

\[ P(X = a) \leq P(a - \epsilon \leq X \leq a + \epsilon) = \int_{a-\epsilon}^{a+\epsilon} f(x) \, dx = F(a + \epsilon) - F(a - \epsilon) \]

- for \( \epsilon \to 0 \), \( P(a - \epsilon \leq X \leq a + \epsilon) \to 0 \), hence \( P(X = a) = 0 \)
- What is the meaning of the density function \( f(x) \)?
- \( f(a) \) is a (relative) measure of how likely is \( X \) will be near \( a \)
\( X \sim U(\alpha, \beta) \)

**Definition.** A continuous random variable has a uniform distribution on the interval \([\alpha, \beta]\) if its probability density function \( f \) is given by \( f(x) = 0 \) if \( x \) is not in \([\alpha, \beta]\) and

\[
f(x) = \frac{1}{\beta - \alpha} \quad \text{for} \quad \alpha \leq x \leq \beta.
\]

We denote this distribution by \( U(\alpha, \beta) \).

- \( F(x) = \int_{-\infty}^{x} f(x) \, dx = \frac{1}{\beta - \alpha} \int_{\alpha}^{x} 1 \, dx = \frac{x - \alpha}{\beta - \alpha} \quad \text{for} \quad \alpha \leq x \leq \beta \)

See R script
For $X \sim \text{Geo}(p)$, we have: 

$F(x) = P(X \leq x) = 1 - (1 - p)^{\lfloor x \rfloor}$

extend to reals: 

$F(X) = P(X \leq x) = 1 - (1 - p)^x = 1 - e^{x \cdot \log(1 - p)} = 1 - e^{-\lambda x}$

for $\lambda = \log\left(\frac{1}{1 - p}\right)$

$f(x) = \frac{dF}{dx}(x) = \lambda e^{-\lambda x}$

**Definition.** A continuous random variable has an exponential distribution with parameter $\lambda$ if its probability density function $f$ is given by $f(x) = 0$ if $x < 0$ and

$f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$.

We denote this distribution by $\text{Exp}(\lambda)$.

$\lambda$ is the rate of events, e.g.,

- $\lambda = \frac{1}{10}$ number of bus arrivals per minute, or $\frac{1}{\lambda} = 10$ minutes to wait for bus arrival
- $P(X > 1) = 1 - P(X \leq 1) = e^{-\lambda} = 0.9048$ probability of waiting more than 1 minute.

Exponential is memoryless: $P(X > s + t | X > s) = e^{-\lambda \cdot (s + t)}/e^{-\lambda \cdot s} = e^{-\lambda \cdot t} = P(X > t)$

See R script and seeing-theory.brown.edu
\( X \sim N(\mu, \sigma^2) \)

**Definition.** A continuous random variable has a normal distribution with parameters \( \mu \) and \( \sigma^2 > 0 \) if its probability density function \( f \) is given by

\[
f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2} \quad \text{for } -\infty < x < \infty.
\]

We denote this distribution by \( N(\mu, \sigma^2) \).

- Also called Gaussian distribution
- Standard Normal/Gaussian is \( N(0, 1) \)
  - \( f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \) sometimes written as \( \phi(x) \)
  - No closed form for \( F(a) = \Phi(a) = \int_{-\infty}^{a} \phi(x) \, dx \)
CCDF of $Z \sim N(0, 1)$

E.g., $P(Z > 1.04) = 0.1492$

See R script
**Quantiles**

**Definition.** Let $X$ be a continuous random variable and let $p$ be a number between 0 and 1. The $p$th quantile or 100$p$th percentile of the distribution of $X$ is the smallest number $q_p$ such that

$$F(q_p) = P(X \leq q_p) = p.$$ 

The median of a distribution is its 50th percentile.

- If $F()$ is strictly increasing, $q_p = F^{-1}(p)$
- E.g., for $\text{Exp}(\lambda)$, $F(a) = 1 - e^{-\lambda a}$, hence $F^{-1}(p) = \frac{1}{\lambda} \log \frac{1}{1-p}$

See R script
Simulation

• Not all problems can be solved with calculus!
• Complex interactions among random variables can be simulated
• Generated random values are called realizations
• Basic issue: how to generate realizations?
  ▶ in R: \texttt{rnorm(5)}, \texttt{rexp(2)}, \texttt{rbinom(\ldots)}, \ldots
• Ok, but how do they work?
• Assumption: we are only given \texttt{runif()}!
• Problem: derive all the other random generators
Simulation: discrete distributions

Bernoulli random variables

Suppose $U$ has a $U(0, 1)$ distribution. To construct a $Ber(p)$ random variable for some $0 < p < 1$, we define

$$X = \begin{cases} 1 & \text{if } U < p, \\ 0 & \text{if } U \geq p \end{cases}$$

so that

$$P(X = 1) = P(U < p) = p,$$
$$P(X = 0) = P(U \geq p) = 1 - p.$$ 

This random variable $X$ has a Bernoulli distribution with parameter $p$.

- For $X_1, \ldots, X_n \sim Ber(p)$ i.i.d., we have: $\sum_{i=1}^{n} X_i \sim Binom(n, p)$

See R script
Simulation: continuous distributions

- $F : \mathbb{R} \to [0, 1]$ and $F^{-1} : [0, 1] \to \mathbb{R}$
  - E.g., $F$ strictly increasing
  - N.B., the textbook notation for $F^{-1}$ is $F^{\text{inv}}$
- For $X \sim U(0, 1)$ and $0 \leq b \leq 1$
  $P(X \leq b) = b$
- then, for $b = F(x)$
  $P(X \leq F(x)) = F(x)$
- and then by inverting
  $P(F^{-1}(x) \leq x) = F(x)$
- In summary:
  $F^{-1}(X) \sim F$ for $X \sim U(0, 1)$

See R script
Common distributions

Relationships among common distributions. Solid lines represent transformations and special cases, dashed lines represent limits. Adapted from Leemis (1986).