Common Continuous Distributions

1. Uniform Distribution

A random variable X is said to be uniform over the interval $[a, b]$ if its density is constant over the interval $[a, b]$. Its CDF is thus 0 if $x < a$, $\frac{x-a}{b-a}$ if $a \leq x \leq b$ and 1 if $x > b$.

2. Exponential Distribution

A random variable X is said to have an exponential distribution with rate λ, denoted as $X \sim \text{Exp}(\lambda)$ if its PDF is $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$. Its CDF is thus $F(x) = \int_{t=\infty}^{x} f(t)dt = \int_{t=0}^{x} \lambda e^{-\lambda t} dt = \left[-e^{-\lambda t}\right]_{t=0}^{x} = 1 - e^{-\lambda x}$.

2.1. Exponential vs Geometric

They are similar, both decay exponentially and both are memoryless. But exponential is continuous whereas geometric is discrete.

3. Gamma Distribution

Gamma(α, λ) distribution with the "shape" parameter $\alpha > 0$ and "rate" parameter $\lambda > 0$ has the PDF: $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$ for $x \geq 0$ where $\Gamma(\alpha) = \int_{z=0}^{\infty} z^{\alpha-1} e^{-z} dz$ is a normalising constant (so that density integrates to 1).

3.2. Remark: $\Gamma(k) = (k-1)!$ for integers $k \geq 1$.

4. Normal Distribution

A random variable X is said to have a normal (Gaussian) distribution with mean μ and a standard deviation σ denoted as $X \sim N(\mu, \sigma^2)$ if its PDF is $f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$

The density curve is bell-shaped and symmetric about its mean μ. A normal distribution with $\mu = 0$ and $\sigma = 1$ is called the standard normal distribution, which we denote $N(0, 1)$.

4.3. PDF & CDF of the SND

The PDF & CDF of the SND $N(0,1)$ are respectively:
PDF: $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ with $-\infty < z < \infty$
CDF: $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$ with $-\infty < z < \infty$

4.4. Remark: The CDF $\Phi(z)$ has no closed-form solution so we use the probability table with which we are all too familiar from AP Stats

5. Beta Distributions

The random variable U is said to have a beta distribution with parameters α, β if its density is given by $f(u) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u^{\alpha-1}(1-u)^{\beta-1}$ for $0 \leq u \leq 1$, denoted as $U \sim \text{Beta}(\alpha, \beta)$.

5.5. Remark

$\text{Beta}(\alpha = 1, \beta = 1)$ is $\text{Uniform}(0,1)$.

Functions/Transformations of a Random Variable

If X is a cont. r.v. with density $f_X(x)$, and $Y = g(x)$, what is the distribution of Y? Generally we solve by first finding the CDF for $Y = g(x)$ then differentiating to find the PDF for Y. Usually this involves identifying the CDF for our r.v. (have this memorised) and then running a chain rule.

6. Remark: Suppose X is a cont. r.v. with the PDF $f_X(x)$. The PDF for $Y = aX + b$ is: $f_Y = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$ if $a \neq 0$.

7. Differentiable & Strictly Monotone Transformations

Suppose f_X is the PDF of X and $g(X)$ is diff. and strictly monotone. Then $Y = g(X)$ is a cont. r.v. with PDF $f_Y(y) = f_X(g^{-1}(y)) \cdot |\frac{d}{dy} g^{-1}(y)|$.

8. Transforming to Uniform

Suppose X is a cont. r.v. with CDF F where:

1. F is strictly increasing on some interval I.
2. $F = 0$ to the left of I, and $F = 1$ to the right of I.
3. I may be a bounded interval or an unbounded interval such as the whole real line.
   then $F^{-1}(u)$ is then well defined for $u \in (0, 1)$ and, interestingly, if we define $Y = F(X)$ then $Y \sim \text{Uniform(0,1)}$ since it has the Uniform CDF. This is a useful trick.

9. The converse: how we generate a r.v. from Uniform with some given CDF

Let F be the CDF with the conditions from the previous section and let $U \sim \text{Uniform}(0, 1)$. What is the distribution of $X = F^{-1}(U)$? Notice $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$ where the last equality holds since U is $\text{Uniform}[0, 1]$. This means that X has CDF equal to F.

Joint Probability Distributions for Discrete R.V.

The joint probability mass function (joint PMF) or simply the joint distribution for discrete r.v. $X_1, X_2, \ldots, X_k$ is defined as
$p(x_1, x_2, \ldots, x_k) = P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)$
$= P(\{X_1 = x_1\} \cap \{X_2 = x_2\} \cap \cdots \cap \{X_k = x_k\})$

10. Properties of joint PMF:

1. $p(x_1, x_2, \ldots, x_k) \geq 0$.
2. Define the probability for an event A as
   $P(A) = P((x_1, x_2, \ldots, x_k) \in A) = \sum_{(x_1, x_2, \ldots, x_k) \in A} p(x_1, x_2, \ldots, x_k)$.
3. If we set $A = \Omega$ (the sample space) in (2), then $P(\Omega) = 1$.