Relationships & Convolution

- **Sum of Binomials:** $X_i \sim \text{Bin}(n_i, p) \implies \sum X_i \sim \text{Bin}(\sum n_i, p)$
- **Sum of Poissons:** $X_i \sim \text{Pois}(\lambda_i) \implies \sum X_i \sim \text{Pois}(\sum \lambda_i)$
- **Sum of Gammas:** $X_i \sim \text{Gamma}(\alpha_i, \beta) \implies \sum X_i \sim \text{Gamma}(\sum \alpha_i, \beta)$
  - Note: Same scale β required.
- **Sum of Chi-Squares:** $X_i \sim \chi^2_{(r_i)} \implies \sum X_i \sim \chi^2_{(\sum r_i)}$
- **Sum of Normals:** $X_i \sim N(\mu_i, \sigma_i^2) \implies \sum a_i X_i \sim N(\sum a_i \mu_i, \sum a_i^2 \sigma_i^2)$
  - Sample Mean: $\bar{X} \sim N(\mu, \sigma^2/n)$

Key Distributions Properties

Binomial (n, p)**:** (# successes in n iid Bernoulli trials)

- **PMF:** $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, for $x = 0, \ldots, n$
- **Mean:** np, **Var:** $np(1-p)$, **MGF:** $((1-p) + pe^t)^n$
- $n \to \inf, p \to 0, np = \lambda : Possion(\lambda)$

Poisson (λ)**:** (count of rare events in time/space)

- **PMF:** $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$, for $x = 0, 1, 2, \ldots$
- **Mean:** λ, **Var:** λ, **MGF:** $e^{\lambda(e^t - 1)}$

Gamma(α, β)**:** (waiting time for α events; sum of α Exp(β))

- **PDF:** $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$, for $x > 0$
- **Mean:** $\alpha\beta$, **Var:** $\alpha\beta^2$, **MGF:** $(1 - \beta t)^{-\alpha}$ for $t < 1/\beta$
- **Gamma-χ^2:** $\chi^2_{(r)} = \text{Gamma}(r/2, 2)$; if $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{2X}{\beta} \sim \chi^2_{(2\alpha)}$
- $\text{Exp}(\beta) = \text{Gamma}(1, \beta)$

Chi-Square (χ^2_r)**:** (sum of r squared standard normals)

- **PDF:** $f(x) = \frac{1}{2^{r/2}\Gamma(r/2)} x^{r/2-1} e^{-x/2}$, for $x > 0$
- **Mean:** r, **Var:** $2r$, **MGF:** $(1 - 2t)^{-r/2}$

Normal $(N(\mu, \sigma^2))$**:** (symmetric bell curve; CLT limit)

- **PDF:** $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$, for $-\infty < x < \infty$
- **Mean:** μ, **Var:** σ^2, **MGF:** $\exp(\mu t + \frac{1}{2}\sigma^2 t^2)$

Discrete Uniform $(1, N)$**:** (all N outcomes equally likely)

- **PMF:** $P(X = x) = \frac{1}{N}$, for $x = 1, 2, \ldots, N$
- **Mean:** $\frac{N+1}{2}$, **Var:** $\frac{(N+1)(N-1)}{12}$

Bernoulli (p)**:** (single binary trial: success/failure)

- **PMF:** $P(X = x) = p^x (1-p)^{1-x}$, for $x = 0, 1$
- **Mean:** p, **Var:** $p(1-p)$, **MGF:** $(1-p) + pe^t$

Geometric (p)**:** (# trials until 1st success)

- **PMF:** $P(X = x) = p(1-p)^{x-1}$, for $x = 1, 2, \ldots$
- **Mean:** $1/p$, **Var:** $(1-p)/p^2$, **MGF:** $\frac{pe^t}{1-(1-p)e^t}$

Negative Binomial (s, p)**:** (# trials until s successes)

- **PMF:** $P(X = x) = \binom{x-1}{s-1} p^s (1-p)^{x-s}$, for $x = s, s+1, \ldots$
- **Mean:** s/p, **Var:** $s(1-p)/p^2$
- $\text{NegBin}(1, p) = \text{Geometric}(p)$; $\lambda = s(1-p), s \to \inf : Possion(\lambda)$

Hypergeometric (N, M, n)**:** (sampling w/o replacement)

- **PMF:** $P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$
- **Mean:** $\frac{nM}{N}$, **Var:** $\frac{N-n}{N-1} \cdot n \cdot \frac{M}{N}(1 - \frac{M}{N})$
- As $N \to \infty$, Hypergeom $\to$ Binomial

Uniform $[a, b]$**:** (all values in $[a, b]$ equally likely)

- **PDF:** $f(x) = \frac{1}{b-a}$, for $a \le x \le b$; **CDF:** $F(x) = \frac{x-a}{b-a}$
- **Mean:** $\frac{a+b}{2}$, **Var:** $\frac{(b-a)^2}{12}$, **MGF:** $\frac{e^{bt} - e^{at}}{t(b-a)}$

Exponential (β)**:** (waiting time; memoryless lifetime)

- **PDF:** $f(x) = \frac{1}{\beta} e^{-x/\beta}$, for $x \ge 0$; **CDF:** $F(x) = 1 - e^{-x/\beta}$
- **Mean:** β, **Var:** β^2, **MGF:** $(1 - \beta t)^{-1}$ for $t < 1/\beta$
- $\text{Exp}(\beta) = \text{Gamma}(1, \beta)$

Weibull (γ, β)**:** (lifetime with aging/wear-out effect)

- **PDF:** $f(x) = \frac{\gamma}{\beta} x^{\gamma-1} e^{-x^\gamma/\beta}$, for $x > 0$; **CDF:** $F(x) = 1 - e^{-x^\gamma/\beta}$
- **Mean:** $\beta^{1/\gamma}\Gamma(1 + 1/\gamma)$
- $\text{Weibull}(\gamma = 1, \beta) = \text{Exp}(\beta)$

Beta (a, b)**:** (natural model for probabilities on $(0, 1)$)

- **PDF:** $f(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}$, for $0 < x < 1$
- **Mean:** $\frac{a}{a+b}$, **Var:** $\frac{ab}{(a+b)^2(a+b+1)}$
- $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$
- $\text{Beta}(1, 1) = \text{Uniform}(0, 1)$

Lognormal (μ, σ^2)**:** (right-skewed; lifetime, income)

- If $\log Y \sim N(\mu, \sigma^2)$, then $Y \sim$ Lognormal
- **PDF:** $f(y) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-(\log y - \mu)^2/(2\sigma^2)}$, for $y > 0$
- **Mean:** $e^{\mu + \sigma^2/2}$, **Var:** $e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2}$, **Median:** e^μ

Cauchy (θ)**:** (heavy-tailed; ratio of two std normals)

- **PDF:** $f(x) = \frac{1}{\pi[1 + (x-\theta)^2]}$
- **Median:** θ
- Moments do not exist
- Standard Cauchy $= t_1$ distribution

Joint & Marginal

- **Discrete:** $p_X(x) = \sum_y p(x,y)$ **Continuous:** $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$
- **Expectation:** $E[g(X,Y)] = \iint g(x,y)f(x,y)dxdy$

Conditional Distributions

- **Discrete:** $p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}$ **Continuous:** $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$
- **Properties:**
  - $E[E(Y|X)] = E(Y)$ (Law of Total Expectation)
  - $Var(Y) = E[Var(Y|X)] + Var(E[Y|X])$
  - $E[XY \mid X] = X \cdot E[Y \mid X]$

Independence $X \perp Y$ iff:

- $f(x,y) = f_X(x)f_Y(y)$
- iff $f(x_1, x_2) \equiv g(x_1)h(x_2)$ for nonneg. functions g, h.
- Range is rectangular (support doesn't depend on each other).
- $M(t_1,t_2) = M_X(t_1)M_Y(t_2)$
- $Cov(X,Y) = 0$ (Necessary but NOT sufficient, unless Bivariate Normal).
- If $E[u(X)], E[v(Y)]$ exist, then $E[u(X)v(Y)] = E[u(X)] \cdot E[v(Y)]$.

MGF for Multivariate RVs

- **Joint MGF:** $M_{X,Y}(t_1,t_2) = E[e^{t_1X+t_2Y}]$
- **Marginal MGF from Joint::** $M_X(t_1) = M_{X,Y}(t_1, 0)$, $M_Y(t_2) = M_{X,Y}(0, t_2)$
- **Product Moments:** $E[X^jY^k] = \frac{\partial^{j+k}}{\partial t_1^j \partial t_2^k} M_{X,Y}(t_1,t_2)\Big|_{t_1=t_2=0}$

Covariance & Correlation

- $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - E[X]E[Y]$
- $Cov(U+V, W) = Cov(U,W) + Cov(V,W)$
- $\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}, \quad -1 \le \rho \le 1$
- $Var(aX+bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X,Y)$

Bivariate Normal

- If $(X,Y) \sim$ BVN, then $X \perp Y \iff \rho = 0$.
- Marginals are Normal.
- Conditionals are Normal.

Bivariate Transformations ($U = g_1(X,Y), V = g_2(X,Y)$)

1. Solve for $x = h_1(u,v), y = h_2(u,v)$.
2. Jacobian $J(u,v) = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$, $|J(u,v)| = |J(x,y)|^{-1}$
3. $f_{U,V}(u,v) = f_{X,Y}(h_1(u,v), h_2(u,v))|J(u,v)|$
4. Find joint support $\mathcal{T}$.

Linear Combinations (Matrix Form)

- If $\mathbf{X} \sim N_n(\mu, \Sigma)$ and $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$: $\mathbf{Y} \sim N_m(\mathbf{A}\mu + \mathbf{b}, \mathbf{A}\Sigma\mathbf{A}^\top)$

Statistics

- Sample Mean: $\bar{X}_n = \frac{1}{n}\sum X_i$
- Sample Variance: $S^2 = \frac{1}{n-1}\sum(X_i - \bar{X}_n)^2$

Sampling from Normal $N(\mu, \sigma^2)$

If $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$:

1. $\bar{X}_n \sim N(\mu, \sigma^2/n)$
2. $\bar{X}_n \perp S^2$ (Independent!)
3. $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$
4. $T = \frac{\bar{X}_n - \mu}{S/\sqrt{n}} \sim t_{n-1}$ (Student's t)

Types of Convergence

1. **In Probability ($X_n \xrightarrow{P} X$):** $\lim_{n\to\infty} P(|X_n - X| < \epsilon) = 1$
   - Implies $X_n \xrightarrow{D} X$.
   - WLLN: $\bar{X}_n \xrightarrow{P} \mu$ (if i.i.d, finite var).
2. **In Distribution ($X_n \xrightarrow{D} X$):** $\lim_{n\to\infty} F_n(x) = F(x)$ at continuity points.
   - CLT: $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$
   - MGF Technique: $M_{X_n}(t) \to M_X(t) \implies X_n \xrightarrow{D} X$.

Key Theorems

- **Chebyshev's Inequality:** For any $k > 0$, $P(|X-\mu| \ge k\sigma) \le \frac{1}{k^2}$ or equivalently $P(|X-\mu| \ge t) \le \frac{\sigma^2}{t^2}$
- **Slutsky's Theorem:** If $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} c$ (const):
  - $X_n + Y_n \xrightarrow{D} X + c$, $X_nY_n \xrightarrow{D} cX$, $X_n/Y_n \xrightarrow{D} X/c$ (if $c \neq 0$)
- **Continuous Mapping:** If $X_n \xrightarrow{P} X$ and g continuous, $g(X_n) \xrightarrow{P} g(X)$. (Also holds for $\xrightarrow{D}$).
- Estimator T_n is **consistent** for θ if $T_n \xrightarrow{P} \theta$.
- Estimator T is **unbiased** for θ if $E[T] = \theta$.

Order Statistics

For i.i.d sample $X_1, \dots, X_n$ with pdf f and cdf F:

- **Minimum** $X_{(1)}$: $f_{min}(x) = n[1-F(x)]^{n-1}f(x)$
- **Maximum** $X_{(n)}$: $f_{max}(x) = n[F(x)]^{n-1}f(x)$

Important Integral/Sum Tricks

- **Gamma Integral:** $\int_0^\infty x^{\alpha-1}e^{-x/\beta}dx = \Gamma(\alpha)\beta^\alpha$
- **Binomial Sum:** $\sum_{x=0}^n \binom{n}{x} a^x b^{n-x} = (a+b)^n$
- **Geometric Series:** $\sum_{k=0}^\infty r^k = \frac{1}{1-r}$ ($|r| < 1$)
- **Exp Expansion:** $e^x = \sum_{n=0}^\infty \frac{x^n}{n!}$