Moments

kth order central moment
$$\sigma_k = \mathbb{E}[(X - m_1)^k]$$

Expectation

$$\mathbb{E}[g(X,Y)] = \sum_y \sum_x g(x,y) f_{X,Y}(x,y)$$

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

If $X \geq 0$ (continuous or discrete)
$$\mathbb{E}[X] = \int_0^{\infty} (1 - F_X(x))\; dx$$

Conditional expectation
$$\mathbb{E}[Y|X=x] = \int y f_{y|x}(y|x) dy$$

$$\mathbb{E}[g(Y)|X=x] = \int g(y) f_{y|x}(y|x) dy$$

Law of iterated expectation
$$\mathbb{E}_X[E_{Y|X}\{g(Y)|X\}] = \mathbb{E}[g(Y)] = \sum_x \mathbb{E}[g(Y)|X=x] f_X(x)$$

$$\mathbb{E}[h(X)g(Y)] = \mathbb{E}[h(X)\mathbb{E}[g(Y)|X]]$$
$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]] = \sum_x \mathbb{E}[Y|X=x] f_X(x)$$

If $X \perp\!\!\!\perp Y$, then
$$\mathbb{E}[g(Y)|X=x] = E[g(Y)]$$
$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathrm{Cov}(X,X)$$
Special Case: Let $Y^* = I(Y=y)$
$$\mathbb{E}[Y^*] = \mathbb{E}[I(Y=y)] = \mathbb{P}(Y=y)$$

Special case cont
$$\mathbb{E}[Y^*] = \sum_x \mathbb{E}[Y^*|X=x] f_X(x) = \sum_x \mathbb{P}(Y=y|X=x) f_X(x) = \mathbb{P}(Y=y)$$

Basics

$$\mathrm{Cov}[X,Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$
$$\rho(X,Y) = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X] \cdot \mathrm{Var}[Y]}} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\sigma_X \sigma_Y}}$$

Expectation Algebra

$$\mathbb{E}[X] = \mathbb{E}[X_+] - \mathbb{E}[X_-] \text{ if } X_+ \geq 0$$
$$\mathbb{E}[|X|] = \mathbb{E}[X_+] + \mathbb{E}[X_-] \text{ if } X_- \geq 0$$
$$\mathbb{E}[a + bX] = a + b\mathbb{E}[X]$$
$$\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$$
If $X \perp\!\!\!\perp Y$, then
$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

Variance Algebra

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + 2\,\mathrm{Cov}[X,Y] + \mathrm{Var}[Y]$$
$$\mathrm{Var}[X-Y] = \mathrm{Var}[X] - 2\,\mathrm{Cov}[X,Y] + \mathrm{Var}[Y]$$
If $X \perp\!\!\!\perp Y$, then
$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$
If $X \perp\!\!\!\perp Y$, then
$$\mathrm{Var}[X-Y] = \mathrm{Var}[X] - \mathrm{Var}[Y]$$
$$\mathrm{Var}[XY] = \mathbb{E}[X^2] \cdot \mathbb{E}[Y^2] - (\mathbb{E}[X] \cdot \mathbb{E}[Y])^2$$
$$\mathrm{Var}[X/Y] = \mathrm{Var}[X \cdot (1/Y)] = \mathrm{Var}[(1/Y) \cdot X]$$
$$\mathrm{Var}[aX + bY] = a^2\,\mathrm{Var}[X] + b^2 Var[Y] + 2ab\,\mathrm{Cov}[X,Y]$$

Correlation

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{Var(Y)}}$$

Calc

Chain rule
$$f(g(x))' = f'(g(x))g'(x)$$

Product rule
$$(u \cdot v)' = u' \cdot v + u \cdot v'$$

Quotient rule
$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

Indeterminate Forms
$\frac{0}{0}, \frac{\pm\infty}{\pm\infty}, \infty - \infty, 0 * \infty, 0^0, 1^\infty, \infty^0$

Determinate Forms
$\infty + \infty = \infty, -\infty - \infty = -\infty, 0^\infty = 0, 0^{-\infty} = \infty, \infty * \infty = \infty$

Series

Exponential
$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

Geometric
$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r} \text{ where } |r| < 1$$

$$\sum_{k=1}^{\infty} kr^{k-1} = \frac{1}{(1-r)^2} \text{ where } |r| < 1$$

$$\sum_{k=2}^{\infty} k(k-1)r^{k-2} = \frac{-2}{(1-r)^3} \text{ where } |r| < 1$$

Arithmetico-geometric
$$\sum_{k=1}^{\infty} k(1-p)^{k-1} = \frac{1}{p^2} = p^{-2}$$

$$\sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} = \frac{2}{p^3}$$

Binomial
$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

$$n(x+y)^{n-1} = \sum_{k=1}^{n} k\binom{n}{k} x^{k-1} y^{n-k}$$

$$n(n-1)(x+y)^{n-2} = \sum_{k=2}^{n} k(k-1)\binom{n}{k} x^{k-2} y^{n-k}$$

Binomial Coefficient

$$\binom{n}{k} = \binom{n}{n-k}$$
$$\binom{n}{k} = \binom{n-1}{k-1}\left(\frac{n}{k}\right) \text{ if } k > 0$$
$$(n+1)\binom{r}{n+1} = r\binom{r-1}{n}$$
$$\binom{n}{r+a}\binom{r+a}{r} = \binom{n}{r}\binom{n-r}{a}$$
$$\frac{1}{k+1}\binom{n}{k} = \binom{n+1}{k+1}\frac{1}{n+1}$$

Distribution Functions

$$\lim_{x\to-\infty} F(x) = 0 \quad \lim_{x\to\infty} F(x) = 1$$
$$F(x) \leq F(y) \text{ if } x < y$$
$$\lim_{h\downarrow 0} F(x+h) = F(x)$$

$$F(x) = \int_{-\infty}^{x} f(t)\, dt \text{ where } f(t) \geq 0 \text{ and } \int_{-\infty}^{\infty} f(t)\, dt = 1$$
$$F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y)$$
$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbb{P}(\boldsymbol{X} \leq \boldsymbol{x}) = \mathbb{P}(X_1 \leq x_1, \ldots, X_k \leq x_k)$$
$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f(t_1, \cdots, t_k)\, dt_1 \cdots dt_k$$
$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b,d) - F(a,d) - F(b,c) + F(a,c)$$
Joint Density
$$F(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \left\{ \int_{-\infty}^{y} f(u,v)\, dv \right\} du$$

Random Variables

$$X : \Omega \to \mathbb{R}$$
$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$$
$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\{\omega : X(\omega) \leq x\})$$
$$\mathbb{P}(X > x) = 1 - F(x)$$
$$\mathbb{P}(x < X \leq y) = F(y) - F(x)$$
$$\mathbb{P}(X < x) = \lim_{n\to\infty} F(x - \frac{1}{n})$$
$$\mathbb{P}(X = x) = \mathbb{P}(X \leq x) - \mathbb{P}(X < x) = F(x) - \lim_{n\to\infty} F(x - \frac{1}{n})$$

PDF

Conditional PDF
$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)\, dy} = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)}\, dv$$

Joint PDF
$$f(x,y) = \frac{\partial^2}{\partial y \partial x} F(x,y)$$

Marginal PDF
$$f_X(x) = \int_{-\infty}^{+\infty} f(x,v)\, dv, \; f_Y(y) = \int_{-\infty}^{+\infty} f(u,y)\, du$$

Convolution

Let $Z = X + Y$
$$f_Z(z) = \sum_x f_X(x) f_{Y|X}(z-x|x) = \sum_y f_Y(y) f_{X|Y}(z-y|y)$$

If $X \perp\!\!\!\perp Y$, then
$$f_Z(z) = \sum_x f_X(x)f_Y(z-x) = \sum_y f_Y(y)f_X(z-y)$$
$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x) \, dx = \int_{-\infty}^{\infty} f_{X,Y}(z-y, y) \, dy$$
If $X \perp\!\!\!\perp Y$, then
$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x) \, dx = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y) \, dy$$

Bivariate Normal

$$\Sigma = \text{Cov}(\boldsymbol{X}, \boldsymbol{X}) = \begin{pmatrix} \sigma_1^2 & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_1, X_2) & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$
$$\text{Cov}(X_1, X_2) = 0 \iff X_1 \perp\!\!\!\perp X_2$$
$$\boldsymbol{U} = \boldsymbol{v} + C\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{v} + C\boldsymbol{\mu}, C\Sigma C)$$
$$X_2|X_1 = x_1 \sim \mathcal{N}(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), \sigma_2^2(1-\rho^2))$$
$$\mathbb{E}[X_2|X_1] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X_1 - \mu_1)$$

Distributions arising from $\mathcal{N}$

Gamma distribution
$$\Gamma\left(\frac{1}{2}, \frac{k}{2}\right) \to \chi_k^2 \to t_k \stackrel{D}{=} \frac{\mathcal{N}(0,1)}{\sqrt{\chi_k^2/k}} \to F_{m,k} \stackrel{D}{=} \frac{\chi_m^2/m}{\chi_k^2/k}.(F_{1,k} \stackrel{D}{=} t_k^2)$$
$$X \sim \Gamma(\lambda, p) \iff \lambda X \sim \Gamma(1, p) \stackrel{c \geq 0}{\Longrightarrow} cX \sim \Gamma(\frac{\lambda}{c}, p)$$
$$X_j \sim \Gamma(\lambda, p_j), j = 1, \ldots, k, X_1, \ldots, X_k \perp\!\!\!\perp \Longrightarrow \sum_{j=1}^{k} X_j \sim \Gamma(\lambda, \sum_{j=1}^{k} p_j)$$
$$\begin{cases} X_1 & \sim \Gamma(\lambda, p) \\ X_2 & \sim \Gamma(\lambda, q) \\ X_1 & \perp\!\!\!\perp X_2 \end{cases} \Longrightarrow \begin{cases} Y_1 = X_1 + X_2 & \sim \Gamma(\lambda, p+q) \\ Y_2 = \frac{X_1}{X_1 + X_2} & \sim \text{Beta}(p, q) \\ Y_1 & \perp\!\!\!\perp Y_2 \end{cases}$$

χ_k^2 distribution (k degrees of freedom)
$$Z_1, \ldots, Z_k \stackrel{\text{iid}}{\sim} \mathcal{N}(0,1) \Longrightarrow V = Z_1^2 + \cdots + Z_k^2 \sim \chi_k^2$$
$$f_V(v) = \frac{v^{(k-2)/2}e^{-v/2}}{2^{k/2}\Gamma(k/2)} \sim \Gamma\left(\frac{1}{2}, \frac{k}{2}\right), v \geq 0$$
$$\mathcal{N}(0,1)^2 \sim \chi_1^2 \stackrel{D}{=} \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$$
$$\chi_2^2 = \Gamma\left(\frac{1}{2}, 1\right) = \text{Exp}\left(\frac{1}{2}\right)$$

t_k distribution (k degrees of freedom)
if
$$Z \sim \mathcal{N}(0,1), \ V \sim \chi_k^2, \ Z \perp\!\!\!\perp V$$
then
$$Q = \frac{Z}{\sqrt{V/k}} \sim t_k$$
$$f_Q(q) = c_1(1 + \frac{q^2}{k})^{-(k+1)/2}, q \in \mathbb{R}$$
where $c_1 > 0$ is a constant
$$\frac{1}{1+q^2} \sim \text{Cauchy}(0,1) \stackrel{D}{=} t_1$$
$$t_k \to \mathcal{N}(0,1) \text{ as } k \to \infty$$

F Distribution
if
$$V \sim \chi_m^2, \ W \sim \chi_k^2, \ V \perp\!\!\!\perp W$$
then
$$S = \frac{V/m}{W/k} \sim F_{m,k}$$
$$f_S(s) = c_2 s^{(m-2)/2}(1 + \frac{m}{k}s)^{-(k+m)/2}, s > 0$$
where $c_2 > 0$ is a constant
Special case: if $m = 1$, then
$$F_{m,k} = F1, k \stackrel{D}{=} t_k^2$$

Estimators

Estimate μ
$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$
Estimate σ^2
$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$
Suppose $\{X_1, \ldots, X_n\} \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, then
$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \sim \mathcal{N}(0,1), \ \overline{X}_n \perp\!\!\!\perp S_n^2, \ \mathbb{E}[S_n^2] = \sigma^2, \ \frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2$$

And
$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \sim \mathcal{N}(0,1), \ \sqrt{n}\frac{\overline{X}_n - \mu}{S_n} \sim t_{n-1}$$

Sampling

Inverse Transform
Let $U \sim \text{Unif}(0,1)$, F be some continuous distribution. Assume $F(x)$ is strictly $\uparrow$ in x, then: $X = F^{-1}(U) \sim F$
Discrete
Let $U \sim \text{Unif}(0,1)$, suppose $X \in \{0, 1, 2, \ldots\} \sim F(x)$. We wish to generate X. For $k = 0, 1, 2, \ldots$, define $Y = k \iff F(k-1) < U \leq F(k)$, then $Y \sim F$
Rejection
Suppose $X \sim F_X$ with a known PDF: f_X. Let $U \sim \text{Unif}(0,1)$, $Z \perp\!\!\!\perp U$, $Z \sim f_Z(z)$, where the pdf is known. $\exists$ a known constant $a > 0$, such that $f_X(z) \leq af_Z(z)\forall z$
Event $E = \{aUf_Z(Z) \leq f_X(Z)\}$. $P(Z \leq x|E) = F_X)(x)$

Generating Functions $-$ PGF

(non-negative integer-valued RVs) $-$ No finiteness guarantee
$$G_X(s) = \mathbb{E}(s^X) = \sum_{j=0}^{\infty} s^j\mathbb{P}(X = j) = \sum_{j=0}^{\infty} s^j f_X(j)$$
Factorial moments
$$\mathbb{E}[X(X-1)\cdots(X-k+1)] = G^{(k)}(1)$$
$$X \perp\!\!\!\perp Y \Longrightarrow G_{X+Y}(s) = G_X(s)G_Y(s)$$
If $X_1, X_2, \ldots$ iid with PGF G_X and $N(\geq 0)$ is a rv w/ PGF G_N, then the sum $S_N = X_1 + \cdots + X_N$ has the PGF:
$$G_{S_N}(s) = G_N(G_X(t))$$
$$G_{X_1, X_2}(s_1, s_2) = \mathbb{E}[s_1^{X_1}s_2^{X_2}]$$
If $X_1 \perp\!\!\!\perp X_2$:
$$G_{X_1, X_2}(s_1, s_2) = G_{X_1}(s_1)G_{X_2}(s_2)$$

Generating Functions $-$ MGF

(General RVs)
$$M_X(t) = \mathbb{E}(e^{tX}) = G_X(e^t) = \sum_{j=0}^{\infty} e^{tj}f_X(j) = \int e^{tx}f_X(x) \, dx$$
$$M_{X_1, X_2}(t_1, t_2) = \mathbb{E}[e^{t_1 X_1}e^{t_2 X_2}] = G_{X_1, X_2}(e^{t_1}, e^{t_2})$$
If $X_1 \perp\!\!\!\perp X_2$:
$$M_{X_1, X_2}(t_1, t_2) = M_{X_1}(t_1)M_{X_2}(t_2)$$
For $S_n = X_1 + \cdots + X_n$, if $\{X_i\}_{i=1}^{n}$ are independent, then:
$$M_{S_n}(t) = \prod_{k=1}^{n} M_{X_k}(t)$$
In addition, if $\{X_i\}_{i=1}^{n}$ have the same distribution as X, then:
$$M_{S_n}(t) = (M_X(t))^n$$
For $Y = a + bX$, if M_X exists:
$$M_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{ta+tbX}] = e^{ta}\mathbb{E}[e^{tbX}] = e^{ta}M_X(tb)$$
If $|M_X(t)| < \infty$ for $t \in (-a, a)$, where $a > 0$ is some constant, then:
$$M_X^{(k)}(0) = \mathbb{E}[X^k]$$

Stein's Lemma

If $g(z)$ is differentiable and $\lim_{z \to -\infty} g(z)\phi(z) = 0$ and $\lim_{z \to +\infty} g(z)\phi(z) = 0$
Then for $Z \sim \mathcal{N}(0,1)$: $\mathbb{E}[Zg(Z)] = \text{Cov}(Z, g(Z)) = \mathbb{E}[g'(Z)]$
Applications
For $Z \sim \mathcal{N}(0,1)$: $\mathbb{E}[Z^{2k}] = (2k-1)!!$ if $\lim_{z \to \pm\infty} m(\mu + \sigma z)\phi(z) = 0$, then:
$$Cov(X, m(X)) = \sigma^2\mathbb{E}[m'(X)]$$

CLT

Suppose $\{X_1, \ldots, X_n\} \stackrel{\text{iid}}{\sim} X$, with $\mu = \mathbb{E}[X]$ and $\sigma^2 = \text{Var}(X) \in (0, \infty)$. Define $Z_n = \sqrt{n}\frac{\overline{X}_n - \mu}{\sigma}$. Then
$$Z_n \stackrel{D}{\longrightarrow} Z \sim \mathcal{N}(0,1)$$

LLN

Suppose $\{X_1, \ldots, X_n\} \stackrel{\text{iid}}{\sim} X$, with $\mu = \mathbb{E}[X]$. Then
$$\overline{X}_n \stackrel{D}{\longrightarrow} \mu$$

MGFs

Gamma: $\Gamma(\lambda, \alpha) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha}, t < \lambda$
Binomial: $(q + pe^t)^n$
Poisson: $\exp\{\lambda(e^t - 1)\}$
Geometric: $\frac{pe^t}{1 - (1-p)e^t}, t < -\ln(1-p)$

EDF

$F(x) = P(X \leq x) = \mathbb{E}[I(X \leq x)]$, $\hat{\mathbb{F}}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x) = \frac{1}{n}N$, where $N \sim \text{Bin}(n, F(x))$. $\mathbb{E}[\hat{\mathbb{F}}_n(x)] = F(x)$, $\text{Var}(\hat{\mathbb{F}}_n(x)) = \frac{1}{n}F(x)(1-F(x))$
LLN: $\hat{\mathbb{F}}_n(x) \stackrel{D}{\to} F(x)$, CLT: If $0 < F(x) < 1$, then $\sqrt{n}\frac{\hat{\mathbb{F}}_n(x) - F(x)}{\sqrt{F(x)(1-F(x))}}$