(Bayes) Suppose that $\Omega = \bigcup_{i=1}^{k} C_i$, where $C_i \cap C_j = \emptyset \ \forall i \neq j$. Then,

$$P(C_j|C) = \frac{P(C \cap C_j)}{P(C)} = \frac{P(C|C_j)P(C_j)}{\sum_{i=1}^{k} P(C|C_i)P(C_i)}$$

(Percentile) Let $0 < \alpha < 1$. A $(100 \cdot \alpha)$th percentile of the random variable X is a value x_α such that $\Pr(X \leq x_\alpha) \leq \alpha$ and $\Pr(X \geq x_\alpha) \geq 1 - \alpha$. A median is the 50th percentile.

(Inequalites) (Markov, Chebyshev, Jensen, respectivley) u : nonnegative and ϕ : concave

$$\Pr[u(X) \geq c] \leq \tfrac{E[u(X)]}{c} \quad \forall c > 0$$
$$\Pr[|X - \mu| > c\sigma] \leq \tfrac{1}{c^2} \quad \text{or} \quad \Pr[|X - \mu| > \epsilon] \leq \tfrac{\sigma^2}{\epsilon^2}$$
$$\phi(E[X]) \leq E[\phi(X)].$$

(conditional) The **conditional p.d.f.** of X_2 given X_1 is defined as

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)},$$

and the **conditional expectation** of $u(X_2)$ given $X_1 = x_1$ is $E[u(X_2)|x_1] = \int u(x_2) f(x_2|x_1) dx_2$
Law of Iterated Expectation : $E[X_2] = E[E[X_2|X_1]]$

(Corr Coeff) Given random variables X and Y, we define the correlation coefficient between X and Y as

$$\rho = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

(transform) Let X be a r.v with pdf f_X with support A. $Y = u(X)$ and u is 1:1 (monotone).
Then the pdf of Y over support $B = \{y|y = u(x)\}$ is

$$f_Y(y) = f_X[u^{-1}(y)] \left| \frac{d}{dy} u^{-1}(y) \right| \quad \text{where} \quad u^{-1}(y) = \begin{pmatrix} V_1(y) \\ V_2(y) \end{pmatrix} \text{ and } \left| \frac{d}{dy} u^{-1}(y) \right| = \left| det \begin{pmatrix} \frac{\partial V_1}{\partial y_1} & \frac{\partial V_1}{\partial y_2} \\ \frac{\partial V_2}{\partial y_1} & \frac{\partial V_2}{\partial y_2} \end{pmatrix} \right|$$

(mgf) The **moment generating function** of a random variable X is defined as

$$M(t) \equiv E[\exp(tX)].$$

It is known that the moment generating function is unique and completely determines the distribution of the random variable: if two random variables have the same moment generating function, they have the same distribution. Given an n-dimensional random vector $X = (X_1, \ldots, x_n)$, we can define its **mgf** as

$$M(t) = M(t_1, \ldots, t_n) \equiv E[\exp(t_1 X_1 + \cdots + t_n X_n)] = E[\exp(t' X)]$$

(independent) Two random variables X_1 and X_2 are **independent** of each other if one of the following holds true.

(1) $f(x_1, x_2) = f_1(x_1) f_2(x_2)$.
(2) $\exists g, h$: nonnegative such that $f(x_1, x_2) = g(x_1) h(x_2)$.
(3) $M(t_1, t_2) = M(t_1, 0) \cdot M(0, t_2)$

(indep') If X_1 and X_2 are independent of each other, we have

(1) $\Pr(a < X_1 < b, c < X_2 < d) = \Pr(a < X_1 < b) \cdot \Pr(c < X_2 < d)$
(2) $u(X_1), v(X_2)$ are indep. for any u, v, thus $E[u(X_1) \cdot v(X_2)] = E[u(X_1)] \cdot E[v(X_2)]$.
(3) $Cov(X_1, X_2) = 0$

(binomial) $X \sim b(n,p)$

$$f(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}, \qquad M(t) = \left(1-p+pe^t\right)^n, \qquad E[X]=np, \quad Var(X)=np(1-p)$$

(Poisson) $X \sim Poisson(\lambda)$

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad M(t) = \exp[\lambda(e^t-1)], \qquad E[X] = Var(X) = \lambda$$

(normal) $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad M(t) = \exp\left[\mu t + \frac{\sigma^2}{2}t^2\right]$$

(gamma) $X \sim \Gamma(\alpha, \beta)$

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-\frac{x}{\beta}}, (x>0), \qquad M(t) = (1-\beta t)^{-\alpha}, \qquad E[X] = \alpha\beta, \qquad Var(X) = \alpha\beta^2$$

where $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}e^{-y}dy$ has properties $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$, $\Gamma(1)=1$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$
and $\alpha = 1$,$\beta = \frac{1}{\lambda}$ implies $f(x) = \lambda e^{-\lambda x}$ (exponential), and $\alpha = \frac{n}{2}$,$\beta = 2$ implies χ^2_n (chi-sqaure with n).

(multi-nor) $X \sim N(\mu, \Sigma)$

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|det\Sigma|}}\exp\left[-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right], \qquad M(t) = \exp\left[\mu' t + \frac{t'\Sigma t}{2}\right]$$

(relations) If $X \sim N\left(\mu, \sigma^2\right)$, then $Y = \frac{X-\mu}{\sigma} \sim N(0,1)$
If $X \sim N(0,1)$, then $X^2 \sim \chi^2(1)$.
If $X_i \overset{iid}{\sim} N(0,1)$, then $\sum_{i=1}^n X_i^2 \sim \chi^2(n)$.
If $X_i \overset{iid}{\sim} N(0,1)$, Let $X = (X_1, \ldots, X_n)'$, then $X \sim N(0, I_n)$.
Suppose $X \sim N(\mu, \Sigma)$. Let $Y = AX \sim N(A\mu, A\Sigma A')$ for some $m \times n$ deterministic matrix A.
Suppose $X = \binom{X_1}{X_2} \sim N(\mu, \Sigma)$, then X_1 and X_2 are independent if and only if $E\left[(X_1-\mu_1)(X_2-\mu_2)'\right] = 0$
If $X \sim N(0, I_n)$, then AX and BX are independent of each other if and only if $AB' = 0$.
If $X \sim N(0, I_n)$, then $X'X \sim \chi^2(n)$

(convergence) A sequence of r.v.'s $Y_1, Y_2, \ldots$ **converges in probability** to Y if $\lim_{n\to\infty}\Pr[|Y_n - Y| \geq \varepsilon] = 0, \forall\varepsilon > 0$.

- $Y_1, Y_2, \ldots$ **converges almost sure (a.s.)** to Y if $P[\omega| \lim_{n\to\infty} Y_n(\omega) = Y(\omega)] = 1$
- $Y_1, Y_2, \ldots$ **converges in distribution** to Y if $\lim_{n\to\infty} F_n(y) = F(y)$ for every continuity point of F, where $F_n(y)$ is the cdf of Y_n and $F(y)$ is the cdf of Y.

(LLN) Let $Y_1, Y_2, \ldots$ denote an i.i.d sequence of random variables with $E[Y_i] = \mu$ and $Var(Y_i) = \sigma^2 < \infty$. We then have $\text{plim}_{n\to\infty}\bar{Y}_n = \mu$

(Cent Limit) Let $X_1, X_2, \ldots$ denote i.i.d random variables with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2$. Then,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i - \mu) = \sqrt{n}\left(\bar{X} - \mu\right) \quad \xrightarrow{d} \quad N(0, \sigma^2) \text{ as } n \to \infty$$

(some more) For a random matrix $\mathbf{X}$, and and a nonstochastic matrix A, we have $Var(A\mathbf{X}) = AVar(\mathbf{X})A'$.
If X_i's are independent with mgf's of $M_i(t)$, then the mgf of $Y = \sum_{i=1}^n a_i X_i$ is equal to $\prod_{i=1}^n M_i(a_i t)$.