Name	Distribution	$E[X]$	Var$[X]$	z-transform
Bernoulli(p), $p \in [0,1]$	$p_X(k) = \begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$	p	$p(1-p)$	$1-p+zp$
Binomial(n,p)	$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$	$(zp+(1-p))^n$
Geometric(p)	$p_X(k) = (1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{zp}{1-z(1-p)}$
Poisson(λ)	$p_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}$	λ	λ	$e^{(z-1)\lambda}$

Name	p.d.f. (f_X)	c.d.f. (F_X)	$E[X]$	Var$[X]$
Uniform(a,b), $a < b$	$\begin{matrix} \frac{1}{b-a} & a \le x \le b, \\ 0 & otherwise \end{matrix}$	$\begin{matrix} 0 & x \le a, \\ \frac{x-a}{b-a} & a \le x \le b, \\ 1 & x \ge b \end{matrix}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential(λ), $0 < \lambda$	$\begin{matrix} \lambda e^{-\lambda x} & 0 \le x, \\ 0 & x < 0 \end{matrix}$	$\begin{matrix} 1-e^{-\lambda x} & 0 \le x, \\ 0 & x < 0 \end{matrix}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Pareto(α), $0 < \alpha < 2$	$\begin{matrix} \alpha x^{-\alpha-1} & 1 \le x, \\ 0 & x < 1 \end{matrix}$	$\begin{matrix} 1-x^{-\alpha} & 1 \le x, \\ 0 & x < 1 \end{matrix}$	$\begin{matrix} \infty & \alpha \le 1, \\ \frac{\alpha}{\alpha-1} & 1 < \alpha \end{matrix}$	∞
Normal(μ, σ^2),	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	GROSS	μ	σ^2

3 axioms (E, F mutually exclusive): 1) $P(E) \ge 0$, 2) $P(E \cup F) = P(E) + P(F)$, 3) $P(\Omega) = 1$

Bayes' Law (discrete): $P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{P(E|F) \cdot P(F)}{P(E)}$ (cont.): $f_X(x|Y=y) = \frac{P(Y=y|X=x) f_X(x)}{P(Y=y)}$

Two events are independent if $P(E|F) = P(E)$ or $P(E \cap F) = P(E)P(F)$, they are conditionally independent if $P(E \cap F|G) = P(E|G)P(F|G)$

If X and Y are independent, then $E[XY] = E[X]E[Y]$.

Memorilessness: $P(X > s+t | X > s) = P(X > t)$.

X and Y are independent if $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$.

Z-transform of X is: $\widehat{X}(z) = E[z^X] = \sum_{i=0}^{\infty} p_X(i) z^i$. $\widehat{X}'(1) = E[X]$, $\widehat{X}''(1) = E[X(X-1)]$.

If X, Y are independent, and $Z = X + Y$, then $\widehat{Z}(z) = \widehat{X}(z) \cdot \widehat{Y}(z)$.

Also $\widehat{X}(z) = p \cdot \widehat{A}(z) + (1-p) \cdot \widehat{B}(z)$ if $X = A$ with probability p and $X = B$ with probability $1-p$.

$\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$

If X, Y independent, then $Var(X+Y) = Var(X) + Var(Y)$.

The k^{th} moment is $E[X^k] = \sum_i p_X(i) i^k$, and the k^{th} central moment is defined as $E[(X - E[X])^k]$.

If $\{X\}_n$ are independent and identically distributed random variables, where N is the random variable indicating how many there are and $S = \sum_{i=1}^{N} X_i$, then $E[S] = E[N]E[X]$, $E[S^2] = E[N]Var(X) + E[N^2](E[X])^2$ and $Var(S) = E[N]Var(X) + Var(N)(E[X])^2$

Yao's: Let $\mathcal{A}$ be a class of deterministic algorithms, and let $\mathcal{I}$ be a class of all inputs. Then, $\min_{A \in \mathcal{A}} E[T_A(I_\tau)] \le \max_{I \in \mathcal{I}} E[T_{A_\sigma}(I),]$ where τ and σ are distributions on $\mathcal{I}$ and $\mathcal{A}$.

Chebyshev: If X has finite expected value μ and variance $\sigma^2 \ne 0$, then, $\forall k > 0$, $P(|X - \mu| \ge \frac{1}{k^2}$.

Markov: If $a > 0$, the $P(|X| \ge a) \le \frac{E[|X|]}{a}$.

If $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = aX + b$, then $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.