Sample space $S :=$ set of all possible outcomes

Event $A :=$ Measurable subset of S

$0 \leq \mathrm{P}(A) \leq 1$
$\mathrm{P}(S) = 1, \mathrm{P}(\varnothing) = 0$
For countable (A_n), $\mathrm{P}(\cup A_n) = \sum \mathrm{P}(A_n)$
$\mathrm{P}(A^c) = 1 - \mathrm{P}(A)$
$\mathrm{P}(A) = \mathrm{P}(A \cap B) + \mathrm{P}(A \cap B^c)$
$A \subseteq B \implies \mathrm{P}(A) \leq \mathrm{P}(B)$
If countable (A_n) partition S,
$\mathrm{P}(B) = \sum \mathrm{P}(A_i \cap B) = \sum \mathrm{P}(A_i)\,\mathrm{P}(B|A_i)$
$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$
$\mathrm{P}(\cup A_n) \leq \sum \mathrm{P}(A_n)$
$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$
$\mathrm{P}(A|B) = \frac{\mathrm{P}(A)}{\mathrm{P}(B)}\,\mathrm{P}(B|A)$
For countable (A_n), independent $:=$ For any subcollection, $\mathrm{P}(\cap A_n) = \prod \mathrm{P}(A_n)$
Define $(A_n) \nearrow A := \cup A_n = A, A_i \subseteq A_{i+1}$
Continuity: If $(A_n) \nearrow A$, then $\lim_{n\to\infty} \mathrm{P}(A_n) = \mathrm{P}(A)$

Random variable $:=$ "any" function from S to $\mathbb{R}$

Distribution of random variable is the collection of all probabilities of events in $\mathbb{R}$

Discrete random variable $X := \sum_{x \in \mathbb{R}} \mathrm{P}(X = x) = 1$
Probability function $p_X(x) := \mathrm{P}(X = x)$
Discrete Law of Total Pr.: $\mathrm{P}(B) = \sum_i p_X(x_i)\,\mathrm{P}(B|X = x_i)$
Continuous random variable $X := \mathrm{P}(X = x) = 0$ for all x
Density function $f : \mathbb{R} \to \mathbb{R} := f \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$
Absolute continuous $:=$ Has density f s.t. $\mathrm{P}(a \leq X \leq b) = \int_a^b f(x)dx$
Cumulative distribution function (cdf) $:= F_X(x) = \mathrm{P}(X \leq x)$
$\lim_{x\to-\infty} F_X(x) = 0, \lim_{x\to\infty} F_X(x) = 1, F_X$ incr., $F_X(x) \in [0,1]$
Discrete: $F_X(x) = \sum_{t \leq x} p_X(t)$
Continuous: $F_X(x) = \int_{-\infty}^{x} f(t)dt$
F always right continuous but left continuous depending on $f(x)$
$F_X(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$, $F_X'(x) = f(x)$

If $Y \sim N(-8, 4)$, then $\mathrm{P}(Y \leq -5) = \mathrm{P}((Y+8)/2 \leq (-5+8)/2) = \Phi(\frac{3}{2})$

Discrete: If $Y(s) = h(X(s))$, then $P(Y = y) = P(h(X) = y) = \sum_{x:h(x)=y} P(X = x)$
Absolute Continuous: If $Y(s) = h(X(s))$ and X has density f_X, then if h differentiable and stric. monot., with inverse, then Y also abs. conti., and $f_Y(y) = f_X(h^{-1}(y))/|h'(h^{-1}(y))|$

Joint cdf $:= F_{X,Y}(x,y) = \mathrm{P}(X \leq x, Y \leq y)$
$\mathrm{P}(a \leq X \leq b, c \leq Y \leq d) = F_{X,Y}(b,d) - F_{X,Y}(a,d) - F_{X,Y}(b,c) + F_{X,Y}(a,c)$
Joint probability (discrete) $:= p_{X,Y}(x,y) = \mathrm{P}(X = x, Y = y)$
Marginal distribution of $X := p_X(x) = \sum_y p_{X,Y}(x,y)$
Joint absolutely continuous $:=$ exists $f_{X,Y}$ s.t. $\mathrm{P}(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f_{X,Y}(x,y)dydx$
Marginal: $\mathrm{P}(a \leq X \leq b) = \int_a^b \int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx$
Conditional distribution: $p_{X|Y}(x|y) = \frac{\mathrm{P}(X=x|Y=y)}{\mathrm{P}(Y=y)}$
Conditional density: $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

Independent: For any events A, B, $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$
$F_{X,Y}(x,y) = F_X(x)F_Y(y)$
Discrete: $p_{X,Y}(x,y) = p_X(x)p_Y(y)$

Expected value $:= \mathrm{E}(X) = \sum_x x p_X(x)$ or $\int_{-\infty}^{\infty} x f_X(x)dx$
$\mathrm{E}(aX = bY) = a\,\mathrm{E}(X) + b\,\mathrm{E}(Y)$
$\mathrm{P}(X \leq Y) = 1 \implies \mathrm{E}(X) \leq \mathrm{E}(Y)$
For independent X, Y, $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dxdy$

Let μ_X be $\mathrm{E}(X)$
Variance $:= \mathrm{Var}(X) = \mathrm{E}((X - \mu_X)^2) = \mathrm{E}(X^2) - 2\mu_X\,\mathrm{E}(X) + \mu_X^2 = \mathrm{E}(X^2) - \mathrm{E}(X)^2$

Standard Deviation $:= \mathrm{Sd}(X) = \sqrt{\mathrm{Var}(X)}$

Covariance $:= \mathrm{Cov}(X,Y) = \mathrm{E}((X-\mu_X)(Y-\mu_Y)) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$
If $\mathrm{Cov}(X,Y) > 0$, then X, Y positively correlated
If $\mathrm{Cov}(X,Y) < 0$, then X, Y negatively correlated
Correlation $:= \mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Sd}(X)\,\mathrm{Sd}(Y)}$
Correlation is only defined for non-constant functions

(X_n) converges in probability to $Y := \forall \epsilon > 0, \mathrm{P}(|X_n - Y| > \epsilon) \to 0$
(X_n) converges almost surely (a.s.) to $Y := \mathrm{P}(X_n \to Y) = 1$

SLLN: For any sequence (X_n) iid each with μ. If $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, then $M_n \to \mu$ a.s.

CLT: Probabilities of Z_n converge to those of $Z \sim N(0,1)$, where $Z_n = \frac{M_n - \mu}{\sigma/\sqrt{n}}$ and M_n is the sample mean of n iid X_i with μ and σ^2
$Z_n = (1/nS_n - \mu)/(\sqrt{\sigma^2/n})$

Now, we know $\lim_{n\to\infty} \mathrm{P}(\frac{1/nS_n - \mu}{\sqrt{\sigma^2/n}} \le z) = \Phi(z)$

1 Distributions

Geometric sum: $\sum_{k=0}^{n} ar^k = a\frac{1-r^{n+1}}{1-r}$

Geometric distribution: Z is the number of misses before the first score

Random Variable	PMF/PDF	Expectation	Variance	Transform
Discrete Uniform over $[a,b]$	$p_X(k) = \begin{cases} \frac{1}{b-a+1}, & \text{if } k = a, a+1, ..., b \\ 0, & \text{otherwise} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)(b-a+2)}{12}$	$\frac{e^{sa}(e^{s(b-a+1)}-1)}{(b-a+1)(e^s-1)}$
Bernoulli with Parameter p	$p_X(k) = \begin{cases} p, & \text{if } k = 1 \\ 1-p, & \text{if } k = 0 \end{cases}$	p	$p(1-p)$	$1 - p + pe^s$
Binomial with Parameters p and n	$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, ..., n$	np	$np(1-p)$	$(1-p+pe^s)^n$
Geometric with Parameter p	$p_X(k) = (1-p)^{k-1}p, \quad k = 1, 2, ...$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^s}{1-(1-p)e^s}$
Poisson with Parameter λ	$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, ...$	λ	λ	$e^{\lambda(e^s-1)}$
Continuous Uniform over $[a,b]$	$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \le x \le b \\ 0, & \text{otherwise} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{sb}-e^{sa}}{s(b-a)}$
Exponential with Parameter λ	$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \ge 0 \\ 0, & \text{otherwise} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - s}, \ (s < \lambda)$
Normal with Parameters μ and $\sigma^2 > 0$	$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{(\frac{\sigma^2 s^2}{2} + \mu s)}$