Discrete Definitions
Sample space: set of possible outcomes (Ω)
Event: a subset of Ω
Elementary event: event with one element in it
Random variables: {X=r}
Exclusive: events are exclusive if the associated sets are disjoint
Exhaustive: events are exhaustive if the union = Ω
Density function: function assigning probability to event
Distinguishable: events are distinguishable if you consider them to be different

Axioms
$P(A) \geq 0$
$P(\Omega) = 1$
If A_1, A_2, A_3 disjoint, $P(A_1 \cap A_2 \cap A_3) = P(A_1)+P(A_2)+P(A_3)$

Empty Set Theorem
$P(\varnothing)=0$, from $P(\Omega) = P(\Omega \cap \varnothing)$ and axioms I and III

Summa. Of Elementary Events Th.
$P(A) = \sum_{i=1}^{n} P(\{s_i\})$ by axiom III

Complem. Event The.
$P(A) = 1 - P(\overline{A})$ from $A \cup \overline{A} = \Omega$ and axiom III

Inclusion-Exclusion Theorem
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
(for exclusive events conforms with axiom III trivially)

Conditional Probability
P(B|A): probability of B given A
$P(B \mid A) = \frac{P(B \cap A)}{P(A)}$
A, B independent if $P(B \mid A) = P(B)$, which further implies $P(A \cap B) = P(A)P(B)$

Bayes' Theorem
$$P(A \mid B) = \frac{P(B \mid A)P(A)}{\sum_k P(B \mid A_k)P(A_k)}$$
Using $P(X = r \mid Y = s) = \frac{p_{r,s}}{\sum_k p_{k,s}}$ and
$P(B \cap A) = P(A \cap B) = P(B \mid A)P(A)$

Binomial Theorem
$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^{n-r} y^r$, ${}^nC_r = \frac{n!}{r!(n-r)!}$

Representations
Tree (conditional probabilities), event tree (event probabilities), grid

Uniform Distribution
$P(X = r) = \frac{1}{n-m+1}$

Binomial Distribution
$P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}$
E(X) = np, V(X) = np(1-p)

Multinomial Distribution
$\frac{n!}{r_1! r_2! \ldots r_k!} p_1^{r_1} p_2^{r_2} .. p_k^{r_k}$,
$\sum_k p_k = 1$, $\sum_k r_k = n$: for entities that can be in k states, r_k of them being in each state (n in total)

Geometric Distribution
$P(X = r) = (1-p)^r p$
$E(X) = \frac{1-p}{p}$, $V(X) = \frac{1-p}{p^2}$

Poisson Distribution
$P(X = r) = \frac{\lambda^r}{r!} e^{-\lambda}$
$E(X) = \lambda$, $V(X) = \lambda$

Expectation
$E(f(X)) = \sum_r f(r) P(X = r)$
E(a)=a, E(aX)=aE(X), E(X+Y)=E(X)+E(Y)

Variance
$\sigma^2 = E((X-\mu)^2) = E(X^2) - (E(X))^2$
$\text{Standard Deviation} = \sigma = \sqrt{Variance}$
$V(a)=0$, $V(aX)=a^2V(X)$, V(X+Y)=V(X)+V(Y) (if independ.)

Covariance
$W(X,Y) = E(XY) - E(X)E(Y)$
If X, Y independent W = 0 clearly
V(X+Y)=V(X)+V(Y)+2W(X,Y)

Correlation Coefficient
$$R = \frac{W(X,Y)}{\sqrt{V(X)V(Y)}}$$
-1: complete negative correlation
0: no correlation (!independence)
+1: complete positive correlation

Generating Functions
$G(\eta) = \sum_{r=0}^{\infty} P(X = r)\eta^r$, $G(\eta) = E(\eta^x)$
$G(1) = 1, G'(1) = E(X)$
$G''(1) + G'(1) - (G'(1))^2 = V(X)$
$G_{X+Y}(\eta) = E(\eta^{X+Y}) = G_X(\eta)G_Y(\eta)$
(if X and Y are independent)
For Binomial $G(\eta) = ((1-p) + p\eta)^n$

Difference Equations
Homogenous: transfer all terms to one side so it is equal to 0, guess $u_n = Aw^n$ and substitute/simplify for the *auxiliary equation*. Find the roots w_1, w_2 of the equation and set $u_n = A_1 w_1^n + A_2 w_2^n$. Then find A_1, A_2 using initial conditions.
Watch for double roots! In this case set $u_n = (A_1 + A_2 n) w^n$
Inhomogenous: deem the right hand side (constant) to be 0 and solve normally, then augment by a f(n) = a, bn or cn^2 such that

substituting it into the original left hand side gives the constant. Now apply initial conditions as usual.

Continuous Definitions

Probability density function: function giving instantaneous probability at any point:

$$P(a \le X < b) = \int_a^b f(x)dx$$

Probability distribution function:

$$F(x) = P(X < x) = \int_{-\infty}^{x} f(t)dt$$

Mode: most probable value
Median: if the median is M then

$$\int_{-\infty}^{M} f(x)dx = \int_{M}^{+\infty} f(x)dx$$

$$E(h(X)) = \int_R h(x)f(x)dx$$

$$V(X) = E((X-\mu)^2)$$

Uniform Distribution

$f(x) = \frac{1}{b-a}$, $E(x) = \frac{b+a}{2}$

Exponential Distribution

$f(x) = \lambda e^{-\lambda x}$, $E(x) = \frac{1}{\lambda}$, $V(x) = \frac{1}{\lambda^2}$

Normal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Standard result: $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$

Distribution function for Normal(0, 1) is $\phi(x)$ (error function)

Central Limit Thm.

For $X_1, X_2, \dots X_n$ IID RVs with mean μ and variance σ^2, let Y be the sum of the RVs. Now as n increases Y tends to $N(n\mu, n\sigma^2)$

Bivariate Distribution

Constructed of two RVs, probability density has the form f(x, y)

$$f_X(x) = \int_{y_{MIN}}^{y_{MAX}} f_{XY}(x,y)dy$$

Independent if

$$f_{XY}(x,y) = f_X(x)f_Y(y)$$

Transforms

$$\int_a^b f(x)dx = \int_{y(a)}^{y(b)} f(x(y))\tfrac{dx}{dy}dy$$

Y=y(X) is a RV with probability density $g(y) = f(x(y))\frac{dx}{dy}$

Constraints on pdfs lead to constraints on transformation function: x(y) and y(x) must be single valued and derivative must be defined (and always +VE or –VE, else multi valued)

Uniform Distributions

These are readily generated by computer, transform to others