|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Basic principle of counting | Event 1: m possible outcomes. Event 2: n possible outcomes & indep from event 1 | | | Total: mn possible outcomes  - Can be generalized to more evts | |
| Permutations (Order matters) | Num of diff arrangements of n objs | | | n! | |
| Permute n objs, of which n1 are same objs, n2 are same objs ... and nr are same objs | | |  | |
| n men sitting in a circle | | | (n-1)! | |
| Combination (Order not impt) | Select m objs from n objs when order not impt(e.g. AB, AC, AD, BC, BD, CD) | | | | = = |
| Combinatorial arg proof. If obj 1 chosen: ways of selecting r-1 objs from remaining n-1 objs. If obj 1 not chosen: ways of selecting r objs from remaining n-1 objs | | | | = + ,  1 ≤ r ≤ n |
| Binomial Theorem. | + + + … = + + + … | | | = 0 |
| Num of subsets of set with n elems | | = 2n (let x = y = 1) | | |
| Multinomial Theorem | Num of divisions of n distinct objs into r **distinct** grps of size n1, n2, ...nr where n1 + n2 +...+nr = n (grp matters) | | = = | | |
| Divide n objs into r grps of m each (grp don't matter)  (e.g. AB CD, AC BD, AD BC) | | = (m! r times) | | |
|  | | | | |
| Num of integer solns of eqns | There are distinct +ve integer-valued vectors (x1, x2, ..., xr) satisfying x1 + ... + xr = n | | | | |
| There are distinct non-negative integer-valued vectors (x1, x2, ..., xr) satisfying x1 + ... + xr = n  OR n identical objs into r distinct grps | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sample Space | Experiment: outcome not predictable with certainty  Sample space: set of all possible outcomes of the experiment | | | | | | | | | Event: subset of sample space  Just draw venn diagram for everything | | | | |
| EF: event containing all outcomes either in E or F or both E & F  EF OR EF: event containing all outcomes that are both in E & F  EC: event containing all outcomes not in E (complement) | | | | | | | | | EF = {x: x E or x F}  EF = { x: x E and x F } OR  EC = {x: x E} | | | | |
| EF: all outcomes in E are in F (subset) | | | | | | | | | If EF and FE, then E = F | | | | |
| Commutative Laws | | | | EF = FE | | | EF = FE | | Associative Laws | | (E F) G = E (F G) | | (EF)G = E(FG) |
| Distributive Laws | | | (E F) G = EG FG | | | EF G = (EG) (FG) | | | DeMorgan's Laws | | = | = | |
| Axioms of Prob | S = sample space. E = event in S  Axiom 1: 0 ≤ P(E) ≤ 1 | | | | | Axiom 2: P(S) = 1  3: For any seq of mutually exclusive events E1, E2,... (i.e. EiEj = when i ≠ j), P = | | | | | | | | |
| Simple Propo-sitions | 1. P() = 0 | | | | | | | | 2. P = (when sample space is finite) | | | | | |
| 3. Strong law of large nums shows converges to P(E) with prob 1  Using axioms, if experiment is repeated many times, by strong law of large numbers, P(E) = proportion which E will occur | | | | | | | | | | | | | |
| 4. If sample space is finite, Axiom 3 becomes P = for mutually exclusive events E1, E2,... | | | | | | | | | | | | | |
| 5. 3 axioms are basic properties of relative frequencies | | | | | | | | 6. Use 3 axioms to check whether given fn P(E) is prob fn | | | | | |
| 7. P(EC) = 1 - P(E) | | | | | | | | 8. If E F, then P(E) ≤ P(F) | | | | | |
| 9. P(EF) = P(E) + P(F) - P(EF) | | | | | | | | 10. P(EFG) = P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG) | | | | | |
| 11. Inclusion-exclusion identity | | | | | | | | P(E1E2...En) = – + ...  + (-1)r+1 + ... + (-1)n+1P(E1E2...En) | | | | | |
| 12i. P ≤ | | | | | | | | 12ii. P ≥ – | | | | | |
| 12iii. P ≤ – + | | | | | | | | | | ... result can be generalized further on | | | |
| Sample spaces with equally likely outcomes | | | If all events in sample space are equally likely to occur, i.e. S = {e1, e2, ..., en}. P({e1}) = P({e­2}) = ... = P({en})  Then for any event E, take P(E) = = . Since P(.) satisfy all 3 axioms, thus P(.) is a probability fn. | | | | | | | | | | | |
| ex = = 1 - 1 + - + - ... | | | | | | | | ex = | | | |
| Prob as a cts set fn | | Incr seq: E1 E2 E3 ... En En+1... Hence =  Decr seq: E1 E2 ... En En+1 ... And = | | | | | | | | If {En, n ≥ 1} is either incr or decr seq, then = P() | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Conditional Prob and Reduced Sample Space | | | If P(F) > 0, P(E|F) = | | | Finding conditional prob using reduced sample space easier | | |
| Multiplication rule: P(AB) = P(A)P(B|A) | | | P(E1E2...En) = P(E1) P(E2|E1)P(E3|E1E2)...P(En|E1E2...En-1) | | |
| Thrm of Total prob and Bayes' Thrm | | Conditioning formula: P(E) = P(E|F)P(F) + P(E|FC)P(FC) | | | | P(F|E) = = | | P(FC|E) = = |
| Thrm of total prob: Suppose F1, F2, ..., Fn are mutually exclusive events s.t. = S, then P(E) = = | | | | | Bayes Thrm: P(Fj|E) = = | |
| Indep Events | E and F are independent P(EF) = P(E)P(F) P(E|F) = P(E) | | | | E and F are independent E and FC are independent | | | |
| E, F and G are indep if P(EF) = P(E)P(F), P(EG) = P(E)P(G), P(FG) = P(F)P(G), P(EFG) = P(E)P(F)P(G) | | | | E, F indep and E, G indep ≠ E, FG indep  If E, F, G are indep, then E will be indep of any event formed from F and G | | | |
| P(E|F) is a prob | Conditional prob satisfies 3 axioms | | | a. 0 ≤ P(E|F) ≤ 1. b. P(S|E) = 1  c. If Ei, i = 1,2,... are mutually exclusive events, then P = . | | | | |
| Note: Results that are true for unconditional prob also true for conditional prob | | | Let Q(E) = P(E|F), then Q(E) is a prob. fn on events of S. Then  E.g. Q(E1 E2) = Q(E1) + Q(E2) - Q(E1E2) P(E1 E2|F) = P(E1|F) + P(E2|F) - P(E1E2|F)  E.g. Q(E1) = Q(E1|E2)Q(E2) + Q(E1|)Q(). So P(E1|F) = P(E1|E2F)P(E2|F) + P(E1|F)P(|F) | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Pmf & Cdf | | For discrete r.v., Prob mass fn (pmf) or prob fn (pf), p(x) = pX(x) =  For r.v., Cumulative dist fn (cdf), F(x) = FX(x) = P(X ≤ x), x | | | | | Properties: 1) P(X = xi) ≥ 0 for i . 2) P(X = xi) = 0 for i .  3) = 1  To check Bin pmf valid, check p + q = 1. = (p+q)n = 1n.  For a discrete r.v., F(a) = | | | | |
| E(x) or | | E(X) = | | | | For a nonnegative integer-valued r.v. Y, E(Y) = = | | | | | |
| E[g(X)] | | E[g(X)] = | | E(aX + b) = aE(X) + b | | nth moment of X, E(Xn) = | | | | | kth central moment = E[] |
| Variance | | Var(X) = E(X - )2 = = E(X2) - [E(X)]2 = | | | | Var(aX + b) = a2Var(X) | | | | | = standard deviation (SD) |
| Formulas | | If X ~ Binomial(n, p) and Y ~ Binomial(n-1, p), then E(Xk) = np \* E[(Y + 1)k-1] | | | | | | | If X ~ Binomial(n,p), P(X = k) = P(X = k-1), k = 1,2,..., n | | |
| Random Variables (r.v.) | |  |  |  |  | | --- | --- | --- | --- | | Bernoulli | X ~ Be (p) | X can only be success (p) or failure (1-p) | P(X = x) = | | Binomial | X ~ Bin(n, p) | k = num of success in n Bernoulli trials | P(X = k) = pk(1-p)n-k, k = 0,1,...,n | | Geometric | X ~ Geo(p) | k = num of Bernoulli trials until success is obtained | P(X = k) = (1-p)k-1p, k = 1,2,... | | Negative binomial | X ~ NB(r, p)  Geo(p) = NB(1, p) | k = num of Bernoulli trials to obtain r successes | P(X = k) = pr(1-p)k-r, k ≥ r | | Hypergeometric | X ~ H(n, N, m) | N distinct balls: m red, N-m blue. Choose n balls w/o replacement. k = num of red balls chosen | P(X = k) = , k = 0,1,...,n | | | | | | | | | | | |
| Poisson r.v. | R.v. X is Poisson with parameter if for some > 0, P(X = k) = , k = 0,1,2...  X~Poisson().  P(X = i+1) = = P(X = i), i = 0,1,...  P(X = 0) = , P(X = 1) = ... | | | | Poisson dist: 1. Approximation to binomial dist w large n and small p  2. Num of events occuring at random at certain points in time | | | | | | |
| 1. If X ~ Binomial(n,p), n is large and p is small, then X~Poisson() approximately, where = np.  Poisson approximation still valid even if trials are not indep, provided their dependence is weak | | | | | | |
| 2. a) Prob only 1 event occuring in interval of length h = h + o(h)  b) Prob ≥ 2 events occuring in interval of length h = o(h)  c) For any integers j1, j2, ... jn and any set of n nonoverlapping intervals, if Ei = event exactly ji  of events occur in the ith intervals, then E1, E2, ... En are indep  Let N(t) = num of events occuring in time interval [0,t]. If above 3 assumptions true, then N(t) ~ Poisson(t), where is rate of occurrences of events per unit time. | | | | | | |
| o(h) | | Notation. Little o h: o(h) stands for any fn f(h) s.t. = 0. E.g. h2 | | | | | | Note. o(h) + o(h) = o(h). Since e.g. h2 + h2 = 2h2 = o(h) | | | |
| Expected Value of Sum of r.v. | | | For a r.v. X, let X(s) denote value of X when s S  E(X) = = , where si = {s: X(s) = xi} | | | | | | | For r.v. X1, X2, ..., Xn, E = | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Intro | X is a Continuous r.v. if a nonnegative fn f, defined x (-∞, ∞), having property that P(X B) = , f(x) = prob density fn (pdf) | | | | | | | | | | | | | | | | | | | | | | | | |
| Properties of pdf | | | | P(X (-∞, ∞)) = = 1 | | | | | | P(a ≤ X ≤ b) = | | | | | P(X = a) = = 0 | | | | dist fn = FX(a) = P(X < a) = P(X ≤ a) = | | | | | f(x) = F(x),  F(x) = cdf |
| Interpretation of pdf at x = f(x) | | | | | | | | | | | | | | P(x < X < x + dx) = ≈ f(x)d(x) (area of rectangle). f(x) ≈ | | | | | | | | | | |
| If X ~ cdf F(x), then F(X) = U ~ uniform(0,1) | | | | | | | | | | | | | | | | | X = F-1(U) ~ cdf F(x) | | | | | | | |
| Expectation and Variance | | | | = E(X) = | | | | | | | | | Lemma. If Y ≥ 0, E(Y) = | | | | | | | If X ~ pdf f(x), E[g(X)] = | | | | | |
| Corollary: E(aX + b) = aE(X) + b | | | | | | | | | | | | | | Var(X) = E[(X - )2] = E(X2) - [E(X)]2 | | | | | | | |
| Uniform r.v. | | | X ~ Uniform (, ) | | | pdf f(x) = | | | | | | | | | | | | E(X) = , Var(X) = | | | | If X ~ Uniform (, ), then  ~ Uniform(0,1) | | | |
| Normal r.v. | | | X ~ N (, ) | | | pdf f(x) = , -∞ < x < ∞ | | | | | | | | | | | | E(X) = , Var(X) = | | | | If X ~ N (, ), then ~ N(0,1) | | | |
| Standard normal dist: Z ~ N (0,1) | | | | F(z) = P(Z ≤ z) = | | | | | | | | | | If Y ~ N(, ) and a is constant, Fy(a) = P(Y ≤ a) = P( ≤ ) = F() | | | | | | | | |
| P(Z ≥ 0) = P(Z ≤ 0) = .5 | | | | -Z ~ N(0, 1) | | | | | P(Z ≤ x) = 1 - P(Z > x) | | | | | P(Z ≤ -x) = P(Z ≥ x) | | | | | aZ + b ~ N(b, a2) | | | |
| The normal approximation to the binomial distribution | | | | | | | | | | | | | | If Sn ~ Binomial(n,p), then ~ N(0,1) approximately for large n | | | | | | | | |
| Exponen-tial r.v. | | | X ~ Exponential() or Exp() for some > 0 | | | | | | | pdf f(x) = | | | | | | | | dist fn, FX(x) = 1 - , x ≥ 0 | | | | | P(X > s) = | | |
| Memoryless Property: P(X > s + t|X > t) = P(X > s) s, t > 0 – (1) | | | | | | | | | | | | | | | P(X > s + t) = P(X > s)P(X > t) s, t > 0 (derived from (1)) | | | | | | | |
| Other cts dist | | X ~ Gamma(, ) where > 0 and > 0 if pdf f(x) = where gamma fn = | | | | | | | | | | | | | | | | | | | | | | | |
| () = ( - 1)( - 1), > 1 | | | | | | | (1/2) = | | | | | | | | If events are occurring randomly in time and follows 3 assumptions of poisson r.v. N(t) ~ Poisson(t), then amt of time one has to wait until total of n events has occurred is a gamma r.v. with parameters (n, ) | | | | | | | | |
| If is an int, say = n, then (n) = (n-1)! | | | | | | | | | | | | | | |
| If X ~ Gamma(1, ), then X ~ Exponential() | | | | | | | | | | | | | | |
| If X ~ Gamma(, ), then E(X) = / and Var(X) = / | | | | | | | | | | | | | | | Xi ~ Exp(), i = 1,2,... and X­i are indep. Then X1 + X2 + ... + Xn ~ Gamma(n, ). Similar to negative binomial for discrete case | | | | | | | | |
| If X ~ Gamma(, ), then X ~ Chi-square dist with n deg of freedom | | | | | | | | | | | | | | |
| X ~ Beta(a,b) if pdf f(x) = where B(a,b) = | | | | | | | | | | | | | | | | | | | E(X) = , Var(X) = | | | | |
| Beta(1,1) = Uniform(0,1) | | | | | | B(a,b) = | | | | | | X ~ Cauchy() with -∞ < < ∞ if pdf f(x) = , -∞ < x < ∞ | | | | | | | | | | E(Xn) DNE for n = 1,2,... | |
| Weibull Dist W(, a, b), fX(x) = | | | | | | | | | | | | | | E(X) = a(1 + ). Var(X) = a2. W(1, , 0) = Exp() | | | | | | | | | |
| Approxi-mation of Binomial r.v. | | | Let X ~ Binomial(n,p). Assume n is large(≥30).  1. Normal approximation. Binomial(n,p) ≈ N(np, npq)  OR ≈ Z, where Z ~ N(0,1). Approximation good if npq ≥ 10  Continuity correction. When finding prob of X using the normal dist  P(X = k) = P(k-1/2 < X < k+1/2). P(X ≥ k) = P(X ≥ k-1/2). P(X ≤ k) = P(X ≤ k+1/2) | | | | | | | | | | | | | | | | 2. Poisson dist. Used when n is large, p is small and np is moderate.  Rule of thumb: use Poisson approximation if p < 0.1 and put = np. If p > 0.9, put = n(1-p) and work in terms of 'failure' | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Joint Dist Fn | | F(x,y) = P(X ≤ x, Y ≤ y) | | | | Note {X > a, Y > b} ≠ {X ≤ a, Y ≤ b}C | | |
| Marginal cdf of X, FX(x) = | | | | Marginal cdf of Y, FY(y) = | | |
| P(X > a, Y > b) = 1 - FX(a) - Fy(b) + F(a,b) | | | | P(a1 ≤ X ≤ a2, b1 ≤ Y ≤ b2) = F(a2, b2) + F(a1, b1) - F(a1, b2) - F(a2, b1) | | |
| If X and Y are discrete r.v. then their joint pmf p(i, j) = P(X = i, Y = j) | | | |  | | |
| Marginal pmf of X, P(X = i) = | | | | Marginal pmf of Y, P(Y = j) = | | |
| If X and Y are cts r.v., then their joint pdf P[(X, Y) C] = = vol under the surface f(x, y) over the region C | | | | | If C = {(x,y): x A, y B}, then P(X A, y B) = | |
| Joint cdf F(a, b) = P(X (-∞, a], Y (-∞, b]) = | | | | f(a, b) = F(a, b) | | |
| Interpretation of joint pdf f(a, b) (density) | | | P(a < X < a + da, b < Y < b + db) = ≈ f(a,b) da db | | | |
| Marginal pdf of X, fX(x) = | | | | Marginal pdf of Y, fY(y) = | | |
| Indep r.v. | | X and Y are indep: P(X A, Y B) = P(X A)P(Y B) | | | | X and Y indep: P(X ≤ a, Y ≤ b) = P(X ≤ a)P(Y ≤ b) OR F(a, b) = FX(a)FY(b) | | |
| Discrete case: X and Y indep: P(X = x, Y = y) = P(X = x)P(Y = y) x,y | | | | Cts case: X and Y indep: f(x, y) = fX(x)fY(y) x,y | | |
| X and Y are indep if knowing value of one does not change dist of other | | | | X and Y indep iff their joint pdf/pmf can be expressed as f(x, y) = g(x)h(y), \*\*-∞ < x < ∞, -∞ < y < ∞\*\* | | |
| Independence is a symmetric relation. If X is indep of Y, then Y is indep of X | | | | | | |
| Sum of indep r.v. | Suppose X and Y are indep cts r.v., then FX+Y(a) = , fX+Y(a) = | | | | | | | |
| Dist of sums of indep r.v. | | 1. Xi ~ Gamma(ti, ), i = 1,...,n ~ Gamma(, ) | | | 3. Zi ~ N(0, 1), i = 1,...,n ~ (chi-square w n deg of freedom)  5. X ~ Poisson(), Y ~ Poisson() X + Y ~ Poisson( + )  6. X ~ Binomial(n, p), Y ~ Binomial(m, p) X + Y ~ Binomial(n + m, p) | | |
| 2. Xi ~ Exp(), i = 1,...,n ~ Gamma(n, ) | | |
| 4. Xi ~ N(, ), i = 1,...,n ~ N(, ) | | |
| Conditional dist for Discrete Case | | If P(F) > 0, P(E|F) = | | conditional pmf of X given Y = y is pX|Y(x|y) = P(X = x|Y = y) = = , y s.t. pY(y) > 0, and p(x,y) is joint pmf of X and Y | | | | |
| If X is indep of Y, then pX|Y(x|y) = px(x), i.e. P(X=x|Y=y) = P(X=x) x,y | | Conditional dist fn of X given Y = y is FX|Y(x,y) = P(X ≤ x|Y = y) = = | | | | |
| Conditional dist for Cts case | | If X and Y have joint pdf f(x,y), then conditional pdf of X given Y = y is f­X|Y(x|y) = , y s.t. fY(y) > 0 | | | | | | |
| Conditional prob/cumulative dist fn of X given Y = y is FX|Y(a|y) = P(X ≤ a|Y=y) = | | | | | | X and Y indep: f­X|Y(x|y) = fX(x) |
| Joint Prob Dist of fns of r.v. | | Let X1 and X2 be jointly cts r.v. w joint pdf (x1, x2). Suppose Y1 = g1(X1, X2) and Y2 = g2(X1, X2) and assume Y1 and Y2 satisfy:  1. x1 and x2 can be uniquely expressed in terms of y1 and y2  2. y1 and y2 have cts partial derivatives at all point (x1, x2) and J(x1, x2) = = – ≠ 0 at all point (x1, x2)  Then Y1 and Y2 are jointly cts w joint pdf (y1, y2) = (x1, x2), where x1, x2 are expressed in terms of y1 y2 | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Expectation of Sums of r.v. | P(a ≤ X ≤ b) a ≤ E(X) ≤ b | | | | | | | If X and Y have a joint pmf p(x,y), then E[g(X, Y)] = | | | | |
| X ≥ Y E(X) ≥ E(Y) | | | | | | | If X and Y have a joint pdf f(x,y), then E[g(X, Y)] = | | | | |
| If E(X) & E(Y) are finite, E(X + Y) = E(X) + E(Y) | | | | | | | If E(Xi) is finite for i = 1,...,n, then E(X1 + ... + Xn) = E(X1) + ... E(Xn) | | | | |
| Covariance, Variance of Sums, Correlations | X and Y indep E[g(X)h(Y)] = E[g(X)]E[h(Y)] | | | | | | | Cov(X, Y) = E[(X - E[X])(Y - E[Y])] (measure dirn of linear relationship btw X and Y) | | | | |
| Cov(X, Y) = E(XY) - E(X)E(Y) | | | | | | | X and Y indep Cov(X, Y) = 0. (opp not necessary; can have non linear r/s) | | | | |
| Cov(X, Y) = Cov(Y , X) | | Cov(X, X) = Var(X) | | | | | Cov(aX, Y) = aCov(X, Y) | | | Cov() = | |
| Var() = + 2 | | | | | | If X1,...,Xn are pairwise indep, i.e. Xi, Xj indp for i≠j, then Var() = | | | | | |
| Correlation of 2 r.v X and Y, (X, Y) = if Var(X)Var(Y) > 0 | | | | | | | | measure strength and dirn of linear r/s | | | -1 ≤ (X, Y) ≤ 1 |
| (X, Y) = 1 Y = a + bX, b = > 0 | | | | (X, Y) = -1 Y = a + bX, b = – < 0 | | | | | | X, Y indep (X, Y) = 0. (converse not true) | |
| E() = | Var() = | | | s2 = | | | | | | E(s2) = | |
| Conditional Expectation | E(X|Y = y) = = , for pY(y) > 0 | | | | | | | | | | E(X|Y = y) = , for fY(y) > 0 | |
| E(g(X)|Y = y) = = | | | | | | | | | | E(g(X)|Y = y) = | |
| E(X) = E(E(X|Y)) (wrt Y, wrt X|Y=y) | | | | E(X) = | | | | | | E(X) = | |
| P(A) = . If Fi = {Y = yi}. Then P(A) = | | | | | | | | | | P(A) = | |
| Conditional Var, Var(X|Y) = E[(X – E(X|Y))2|Y] | | | | | Var(X|Y) = E(X2|Y) - [E(X|Y)]2 | | | | | Var(X) = E[Var(X|Y)] + Var[E(X|Y)] | |
| Moment Generating Functions | Moment Generating Fn: M(t) = E(etX) | | | X is discrete w pmf p(x), M(t) = | | | | | | | Y is cts w pdf f(x): M(t) = | |
| Mn(t) = E(XnetX), n ≥ 1 | | | | Mn(0) = E(Xn), n ≥ 1 | | | | | mgf unique to each distribution, same as pdf and cdf | | |
| X and Y indep MX+Y(t) = Mx(t)MY(t) | | | |  | | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Markov's Inequality | | | If X is a r.v. that takes only nonnegative values, then for any a > 0, P(X ≥ a) ≤ | | | |
| Chebyshev's Inequality | | | If X is a r.v. w finite mean and var , then for any value of k > 0, P(|X-| ≥ k) ≤ | | | If Var(X) = 0, then P(X = E[X]) = 1 |
| Weak Law of large numbers | | | Let X1, X2,... be a seq of indep and identically distributed r.v. each having finite mean E[Xi] = . Then for any > 0, P 0 as n ∞ | | | |
| Central Limit theorem | Let X1, X2,... be a seq of indep and identically distributed r.v. each having mean and var . Then N(0,1) as n ∞.  i.e. for - ∞ < a < ∞, P( ≤ a) as n ∞. Note = = =  Continuity correction. When using normal dist to approx P(X = k) = P(k-1/2 < X < k+1/2).  P(X ≥ k) = P(X ≥ k-1/2). P(X > k) = P(X ≥ k + 1/2). P(X ≤ k) = P(X ≤ k+1/2) P(X < k) = P(X ≤ k - 1/2) | | | | | |
|  | Lemma. Let Z1, Z2, … be a seq of r.v. having distribution fns mgf , n ≥ 1 and let Z be a r.v. w dist fn FZ and mgf MZ. If (t) MZ(t) for all t, then (t) FZ(t) for all t at which FZ(t) is cts | | | | | |
| Strong Law of large nums | Let X1, X2,... be a seq of indep and identically distributed r.v. each having finite mean E[Xi] = . Then w prob 1, as n ∞ | | | | | |
| One-sided Chebyshev's Inequality | | | | X is r.v. w mean 0, var , P(X ≥ a) ≤ | | |
| Chernoff bounds | | P(X ≥ a) ≤ e-taM(t) for all t > 0. | | | P(X ≤ a) ≤ e-taM(t) for all t < 0 | |
| Jensen's Inequality | | A twice-differentiable real-valued fn f(x) is convex if f''(x) ≥ 0 for all x; concave = f''(x) , 0  If f(x) is a convex fn, then E[f(X)] ≥ f(E(X)), if E(X) exists and is finite | | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Pmf p(x), pdf f(x) | Mgf M(t) | Mean | Variance |
| Bernoulli |  |  | p | p(1-p) |
| Binomial w param n,p; 0 ≤ p ≤ 1 | px(1-p)n-x, x = 0,1,...,n | (pet + 1 - p)n | np | np(1-p) |
| Poisson w param > 0 | , x = 0,1,2... |  |  |  |
| Geometric w param p; 0 ≤ p ≤ 1 | p(1-p)x-1, x = 1,2,... |  |  |  |
| Negative binomial w param r, p; 0 ≤ p ≤ 1 | pr(1-p)x-r, n = r, r+1, ... |  |  |  |
| Hypergeometric | P(X = k) = , k = 0,1,...,n |  |  |  |
| Discrete uniform (a,b) |  |  |  |  |
| Uniform over (a, b) |  |  |  |  |
| Exponential w param > 0 |  |  |  |  |
| Gamma w param (s, ), > 0 |  |  |  |  |
| Normal w param (, ) | , -∞ < x < ∞ |  |  |  |