Permutations: $n!$ ways to permute n items, combination: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ ways to combine k items chosen from n

Geometric series: $\sum_{i=0}^{n-1} ar^i = \frac{a(r^n-1)}{r-1} = \frac{a(1-r^n)}{1-r}$, Harmonic series: $\sum_{i=1}^{n} \frac{1}{i} \approx \ln(n)$ (approximation)

A useful approximation: $e^x \geq 1+x, \forall x$; $e^x \approx 1+x$ if $|x|$ is small

union bound: $P(A_1 \cup \cdots \cup A_n) \leq P(A_1) + \cdots P(A_n)$

throw m balls into n bins (at random), { birthday paradox: $m \geq \sqrt{n}$ for $\geq 50\%$ chance that 2 balls end up in the same bin; coupon collector: $m \geq n\log(n)$ for balls to end up in every single bin

if $m=n$, largest bin $\approx \log(n)$ size

conditional probability: $P(A|B) = \frac{P(A\cap B)}{P(B)}$, Bayes Rule: $P(A|B) = P(A)\frac{P(B|A)}{P(B)}$, independent if $P(A\cap B) = P(A)P(B)$

Expected value: $E(X) = \sum x P(X=x)$ $\forall$ possible values x, OR $\int x p(x) dx$ if continuous

Median of random variable: $P(X \geq m) \geq \frac{1}{2}$ and $P(X \leq m) \leq \frac{1}{2}$

variance: $E(X-E(X))^2 = E(X^2) - E(X)^2$, std dev: $std(X) = \sqrt{var(X)}$,

$\boxed{\text{if } Y = aX+b,\ E(Y) = aE(X)+b,\ var(Y) = a^2 var(X)}$

biased coin with $P(head) = p$, then ① expected $\frac{1}{p}$ toss to get 1st head, ② $E(\text{single toss}) = p$, $var = p(1-p)$

Markov Inequality: if X is a positive RV, then $\forall a > 0$, $P(X \geq a) \leq E(X)/a$

Linearity of Expectation: $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$ whether they are independent or not

variance: $var(X_1 + \cdots + X_n) = var(X_1) + \cdots + var(X_n)$ only if they are independent

Chebyshev Inequality: if RV X has std dev σ, then $\forall k > 0 \in \mathbb{R}$, $P(|X - E(X)| \geq k\sigma) \leq \frac{1}{k^2}$

$E(\text{sum of independent RVs}) = X_1 + \cdots + X_n$, $E(\text{avg of indep RVs}) = \frac{1}{n}(X_1 + \cdots + X_n)$

Hoeffding's Inequality: if $X_1, \cdots, X_n \in [0,1]$ are indpt and $E(X_i) = \mu$, then $P(|(\frac{1}{n})(X_1 + \cdots + X_n) - \mu| \geq \varepsilon) \leq 2e^{-2\varepsilon^2 n}$

Hash table with chaining: $T(\text{bucket}(x))$ lookup (collision = flat to linked list)

Power of 2 choices: $T(\log(\log n))$ lookup (2 hash func, put in less full bucket)

Two-level Hashing: high chance of $O(n)$ lookup (each bucket is another hash table), primary: $O(n)$, secondary: $O(n_1^2 + \cdots + n_n^2)$

Bloom Filter: store n items in $\approx 10n$ bits, initialize all bits to 0, insert: set all hash funcs' result of input to true, and ~~retu~~ lookup will return $T[h_1(x)]$ && $\cdots$ && $T[h_k(x)]$. 10 bits = $P(err) \leq 1\%$

Jaccard Similarity: for 2 sets $A, B \subset U$, $J(A,B) = \frac{|A\cap B|}{|A\cup B|}$

min-hash to estimate Jaccard: pick a random permutation π of U, $h_\pi(A)$ = element of A appears earliest under perm π

Distribution	Domain	Param	Mean	Var	f(density)
Normal	$\mathbb{R}$	μ, σ^2	μ	σ^2	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$
Poisson	$\mathbb{Z}^+$	λ	λ	λ	$e^{-\lambda}\frac{\lambda^k}{k!}$
Binomial	$0 \sim n$	n, p	np	$np(1-p)$	$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$
Multinomial	$\{0, \cdots, n\}^k$	$n, (p_1, \cdots, p_k)$	np_i		

Max likelihood of normal:

$P(data|\theta)$: $P(x_1, \cdots, x_n|\theta) = P_\theta(x_1) \cdots P_\theta(x_n)$

take natural log for log-likelihood

$LL(\theta) = \sum_{i=1}^{n} \ln P_\theta(x_i)$

find θ that maxes $LL(\theta)$ using calc

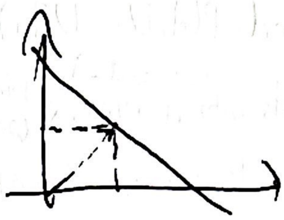
set ~~LL(θ)~~ $LL'(\theta) = 0$

Laplace smoothing: if k observations (eg. binomial $k=2$), add ~~α~~ 1 to each data, then do MLE

covariance: $cov(X,Y) = E(XY) - E(X)E(Y)$, correlation: $corr(X,Y) = \frac{cov(X,Y)}{std(X)std(Y)}$

X, Y independent $\Rightarrow cov(X,Y) = corr(X,Y) = 0$; $corr(X,Y) \neq 0 \Rightarrow$ dependent

test



relationship between density & ~~CPF~~ CDF

point (X_1, X_2): $\{X_1, X_2 \mid 0 \le X_1 \le X_2 \le 1\}$, $Z = max(X_1, X_2)$