Counting

Ordered Sampling with replacement	n^k
Ordered Sampling without replacement	${}^nP_k = \frac{n!}{(n-k)!}$
Unordered Sampling without replacement	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$
Unordered Sampling with replacement	$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}$
The Binomial Theorem	$(x+y)^n = \sum \binom{n}{k} x^k y^{n-k}$
The Multinomial Theorem	$(x_1 + x_2 + ... + x_r)^n = \sum\limits_{\substack{(n_1,n_2,...n_r): \\ n_1+...+n_r=n}} \frac{n!}{n_1!n_2!...n_r!} x_1^{n_1} x_2^{n_2} ... x_r^{n_r}$

There are $\binom{n+k-1}{k-1}$ distinct nonnegative integer-valued vectors $< x_1, ..., x_r >$ satisfying $x_1 + ... + x_k = n$

Probability

Probability Rules

$(\bigcup_{i=1}^n E_i)^c = \bigcap_{i=1}^n E_i$	$(\bigcap_{i=1}^n E_i)^c = \bigcup_{i=1}^n E_i$	$P(E^c) = 1 - P(E)$	$P(E \cup F) = P(E) + P(F) - P(EF)$	$P(E\|F) = \frac{P(EF)}{P(F)}$

$P(E) = P(EF) + P(EF^c) = P(F)P(E|F) + P(F^c)P(E|F^c) = P(F)P(E|F) + (1 - P(F))P(E|F^c)$

Law of Total Probability	$P(E) = \sum_{i=1}^n P(E\|F_i)P(F_i)$
Bayes' Theorem	Suppose that $F_1, F_2, ...F_n$ are mutually exclusive events such that $\bigcup_{i=1}^n F_i = S$ (Sample Space), then $P(F_j\|E) = \frac{P(EF_j)}{P(E)} = \frac{P(E\|F_j)P(F_j)}{\sum_{i=1}^n P(E\|F_i)P(F_i)}$
Independent Events	Two events are independent if $P(EF) = P(E)P(F)$

Discrete Random Variable

Probability Mass Function (PMF)	$p_X(a) = P\{X = a\}$
Cumulative Distribution Function (CDF) of X	$F_X(a) = P\{X \le a\} = \sum_{x \le a} p_X(a)$
Expected Value	$E[X] = \sum_x x p_X(x)$
For any function of g	$E[g(X)] = \sum_x g(x) p_X(x)$
$Var(X) = E[X^2] - (E[X])^2$	$SD(X) = \sqrt{Var(X)}$ $Var(aX + b) = a^2 Var(X)$

Name	PMF	Mean	Variance
$Bernoulli(p)$	$P\{X = 1\} = p,\ P\{X = 0\} = 1 - p$	p	$p(1-p)$
$Binomial(n, p)$	$P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, ..., n$	np	$np(1-p)$
$Geometric(p)$	$P\{X = n\} = p(1-p)^{n-1}, n = 0, 1, 2, ...$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$Poisson(\lambda)$	$P\{X = n\} = e^{\lambda} \frac{\lambda^n}{n!}, n = 0, 1, ...$	λ	λ
$NegativeBinomial(r, p)$	$P\{X = r\} = \binom{n-1}{r-1} p^r (1-p)^{n-r}, n = r, r+1, ...$	$\frac{r}{p}$	$r\frac{(1-p)}{p^2}$
$hypergeometric$	$P\{X = i\} = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}, i = 0, 1, 2, ..., min(n, m)$	$\frac{nm}{N}$	$\frac{N-n}{N-1} np(1-p) where p = m/N$

Continuous Random Variables

Probability Mass Function (PMF)	$f_X(a) = \frac{1}{\epsilon} P\{a - \frac{\epsilon}{2} <= X \le a + \frac{\epsilon}{2}\}$
Cumulative Distribution Function (CDF) of X	$F_X(a) = P\{X \le a\} = \int_{-\infty}^a p_X(a) dx$
$P\{X \in B\} = \int B f_X(x) dx$	$f_X(x) = \frac{d}{dx} F_X(x)$
Expected Value	$E[X] = \int_{-\infty}^{\infty} x f_X(x)$
For any function of g	$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)$
$Var(X) = E[X^2] - (E[X])^2$	$SD(X) = \sqrt{Var(X)}$ $Var(aX + b) = a^2 Var(X)$

Name	PMF	CDF	Mean	Variance
$uniform(a, b)$	$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{othewrise} \end{cases}$	$F(x) = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & \text{if } a < x < b \\ 1 & x \ge a \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$Exponential(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \ge 0 \\ 0 & \text{othewrise} \end{cases}$	$F(x) = 1 - e^{-\lambda x}$ if $a \ge 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$Normal(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$	$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$ *Use normal table	μ	σ^2

For any normal random variable X with parameters (μ, σ^2), $Z = \frac{X-\mu}{\sigma}$ is the standardized normal random variable.
DeMoivere-Laplace Limit Theorem: If S_n denotes the number of successes that occur when n independent trials, each resulting in asuccess with probability p, are performed then, for any a ¡ b, $P\{a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\} -> \phi(b) - \phi(a)$
Memoryless RV: We say that a nonnegative random variable X is *memoryless* if $P\{X > s+t | X > t\} = P\{X > s\}$
Distribution of a Function of a RV: $f_Y(y) = f_X[g^{-1}(y)] | \frac{d}{dy} g^{-1}(y)|$ if $y = g(x)$ for some x

Jointly Distributed Random Variables

Note: Replace summation with integration for continuous

Joint CDF	$f_{XY}(a,b) = P\{X=a, Y=b\}$	Joint CDF	$F_{XY}(a,b) = P\{X \leq a, Y \leq b\}$
Marginal PMF of X	$p_X(a) = \sum_b p_{XY}(a,b)$	Marginal PMF of Y	$p_Y(b) = \sum_a p_{XY}(a,b)$
Marginal CDF of X	$F_{XY}(a) = F_{XY}(a, \infty)$	Marginal CDF of Y	$F_X Y(b) = F_{XY}(\infty, b)$

Conditional PMF	$p_{X\|Y}(x\|y) = P\{X = x \| Y = y\} = \frac{P\{X=x, Y=y\}}{P\{Y=y\}} = \frac{p_{XY}(x,y)}{p_Y(y)}$
Conditional CDF	$F_{X\|Y}(x\|y) = P\{X \leq x \| Y \leq y\} = \sum_{a \leq x} p_{X\|Y}(a\|y)$
Given Function $g(X,Y)$	$\sum_y \sum_x g(x,y) p_{XY}(x,y)$

Conditional Jointly Distributed RV

$F_{X\|A}(x) = \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)}$ if $a \leq x < b$	$f_{X\|A}(x) = \frac{f_X(x)}{P(A)}$	$f_{X\|Y}(x,y) = \frac{f_{XY}(x,y)}{f_Y(y)}$
$E[X\|A] = \int_{-\infty}^{\infty} x f_{X\|A}(x) dx$	$E[g(X)\|A] = \int_{-\infty}^{\infty} g(x) f_{X\|A}(x) dx$	

Independent Random Variables: Two random variables X and Y are said to be independent if for every a and b we have $f_{XY}(a,b) = f_X(a) f_Y(b)$.

Summation of Independent Random Variables

CMF	$F_{X+Y}(a) = P\{X + Y <= a\} = \int_{-\infty}^{\infty} F_X(a-y) f_Y(y) dy$
PDF	$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy$

Sum of Normal RV: Mean and Variance are the summation of all the normal RV's mean and variance.
Jacobian Transformation (Joint Probability Distribution functions of RV)

$f_{Y_1,Y_2}(y_1,y_2) = \frac{f_{X_1,X_2}(x_1,x_2)}{\|J(x_1,x_2)\|}$	$J(x_1,x_2) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} = \frac{\partial g_1}{\partial x_1}\frac{\partial g_2}{\partial x_2} - \frac{\partial g_1}{\partial x_2}\frac{\partial g_2}{\partial x_1} \neq 0$	$y_1 = g_1(x_1,x_2)$ and $y_2 = g_2(x_1,x_2)$

Properties of Expectation

Joint PMF	$E[g(X,Y)] = \sum_y \sum_x g(x,y)p(x,y)$	Joint PDF	$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)p(x,y)dxdyy$

$E[aX+b] = aE[X]+b$	$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$ if X and Y are independent

$Cov(X,Y) = E[XY] - E[X]E[Y]$	$Cov(X,Y) = Cov(Y,X)$
$Cov(X,X) = Cov(X,X)$	$Cov(aX,bY) = abCov(X,Y)$
$Cov(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j) = \sum_{i=1}^n \sum_{j=1}^m Cov(X_i, Y_j)$	$Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i) + \sum_{i=1}^n \sum_{j \neq i} Cov(X_i, X_j)$

$Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i)$ if $X1,...,Xn$ are pairwise independent
Correlation: $\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$ where $-1 \leq \rho(X,Y) \leq 1$

Conditional Expectation	$E[X\|Y=y] = \sum_x x p_{X\|Y}(x,y)$	$E[g(X)\|Y=y] = \sum_x g(x) p_{X\|Y}(x,y)$

Conditioning Property	$E[X] = E[E[X\|Y]]$	Expectation of X	$\sum_y E[X\|Y=y]P\{Y=y\}$

Conditional Variance	$Var(X\|Y) = E[X^2\|Y] - (E[X\|Y])^2$
Unconditional Variance of X given Y	$E[Var(X\|Y)] = E[X^2] - E[(E[X\|Y])^2]$
$Var(E[X\|Y]) = E[(E[X\|Y])^2] - (E[X])^2$	$Var(X) = E[Var(X\|Y)] + Var(E[X\|Y])$

General Form Predictor	$E[(Y - g(X))^2] \geq E[(Y - E[Y\|X])^2]$	best case is $g(X) = E[Y\|X]$
Linear Predictor	$E[(Y - (a+bX))^2] \geq E[(Y - E[Y\|X])^2]$	$b = \rho \frac{\sigma_Y}{\sigma_X}$ and $a = E[Y] - bE[X]$

MGF of PDF	$M(t) = E[e^{tX}] = \sum_x e^{tx} p(x)$
MGF of PMF	$M(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx$
More than 2 RV MGF	$M(t_1,...,t_n) = E[e^{t_1 x_1 + ... + t_n X_n}]$
Individual MGF	$M_{X_i}(t) = M(0,...,0,t,0,...,0)$
Joint MGF	$M_{XY}(t,t') = E[e^{tX + T'Y}] = \sum_y \sum x e^{tx+t'y} p_{XY}(x,y)$
MGF of IRV	$M_{X+Y}(t) = M_X(t) M_Y(t)$

Nth Moment: $M^n(0) = E[X^n]$ $n \geq 0$

Moment Generating Functions for Some RV

Binomial	$M(t) = (pe^t + 1 - p)^n$
Poisson	$M(t) = exp\{\lambda(e^t - 1)\}$
Exponential	$M(t) = \frac{\lambda}{\lambda - t}$ for $t < \lambda$
Standard Normal	$M(t) = e^{t^2/2}$
Normal	$M(t) = exp\{\frac{\sigma^2 t^2}{2} + \mu t\}$

Limit Theorems

Markov's Inequality:
If X is a random variable that takes only nonnegative values, then for any value $a > 0$, $P\{X \geq a|\} \leq \frac{E[X]}{a}$
Chebyshev's Inequality:
If X is a random variable with finite mean μ and variance σ^2, then for any value $k > 0$, $P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$
The Weak Law of Large Numbers
Let $X_1, X_2, ...$ be a sequence of iid Random variables, each having finite mean. Then, for any $\epsilon > 0$, $P\{|\bar{X} - \mu| \geq \epsilon\} \to 0$ where $\bar{X} = \frac{X_1 + X_2 + ... + X_n}{n}$ as $n \to \infty$
The Central Limit Theorem
Let $X_1, X_2, ...$ be a sequence of iid random variables, each having mean and variance. Then, the distribution of $Z = \frac{X_1 + ... + X_n - n\mu}{\sigma\sqrt{n}}$ tends to the standard normal as $n \to \infty$. That is, for $-\infty < a < \infty$, $P\{Z \leq a\} = \phi(a) \to N(0,1)$ as $n \to \infty$

Random Processes

Definitions
Random Processes: a collection of random variables usually indexed by time
Sample Function: the time function $x(t, s)$ associated with the outcome of s of an experiment
Ensemble: the set of all possible time function that can result from an experiment
Random Sequence: A random sequence X_n is an ordered sequence of random variables $X_0, X_1, ...$

Processes
Bernoulli Process: $X(t)$ is a sequence of Bernoulli trials; trials are independent of each other
$P\{X_n = 1\} = p = 1 - P\{X_n = 0\}$
Counting Process Given a stochiastic process $N(t)$

$N(0) = 0$	$N(t) \in \{0, 1, 2, ...\}$ for all $t \in [0, \infty)$	for $0 < s < t$, $N(t) - N(s)$ shows the no. of events in $(s, t]$.

Poisson Process Given $\lambda > 0$, A counting process $N(t), t \in [0, \infty)$ is called a Poisson Process with rate λ if the following conditions hold:

$\to N(0) = 0$
$\to N(t)$ has independent and stationary increments
$\to$ The number of arrivals in any interval of length $\tau > 0$ has $poisson(\lambda\tau)$ distribution

PMF and Joint PMF

PMF of $M = N(t_1) - N(t_0)$	$p_M(m) = \begin{cases} \frac{[\lambda(t_1 - t_0)]^m}{m!} e^{-\lambda(t_1 - t_0} & m = 0, 1... \\ 0 & \text{otherwise} \end{cases}$
Joint PMF of $N(t_1), ..., N(t_k)$, $t_1 < t_2 < ... < t_k$	$p_{N(t_1),...,N(t_k)}(n_1, .., n_k) = \begin{cases} \frac{\alpha_1^{n_1} e^{-\alpha_1}}{n_1!} \frac{\alpha_1^{n_2 - n_1} e^{-\alpha_2}}{(n_2 - n_1)!} ... \frac{\alpha_k^{n_k - n_{k-1}} e^{-\alpha_k}}{(n_k - n_{k-1})!} & m = 0, 1..., \alpha_i = \lambda(t_i - t_{i-1}). \\ 0 & \text{otherwise} \end{cases}$

Theorem: For a Poisson process of rate λ, the inter-arrival times $X_1, X_2, ...$ are an iid random sequence with the exponential PDF.
Memoryless Property of the Poisson Process: $P\{X_n - x' > x | X_n > x'\} = P\{X_n > x\} = e^{-\lambda x}$
Theorem: A counting process with independent exponential inter-arrivals $X_1, X_2, ...$ with mean $E[X_i] = 1/\lambda$ is a Poisson process of rate λ
Brownian Motion Process: A continuous time, continuous value process. Has the property that $X(0) = 0$ and for $\tau > 0, X(t + \tau) - X(t)$ is a Gaussian random variable with mean 0 and variance $\alpha\tau$ that is independent of $X(t')$ for all $t' \leq t$

Brownian Motion	$X(t + \delta) = X(t) + [X(t + \delta) - X(t)]$
PDF of Y_δ	$P_{Y_\delta}(y) = \frac{1}{\sqrt{2\pi\alpha\delta}} e^{-y^2/2\alpha\delta}$
Joint PDF	$f_{X(t_1),...,X(t_k)}(x_1, ..., x_k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi\alpha(t_n - t_{n_1})}} e^{-(x_n - x_{n_1})^2/2\alpha(t_n - t_{n_1})}$

	Random Process	Random Sequence
Expected Value of a process	$\mu_X(t) = E[X(t)]$	
Autocovariance	$C_X(t, \tau) = Cov(X(t), X(t + \tau))$	$C_X[m, k] = Cov[X_m, X_{m_k}]$
Autocorrelation	$R_X(t, \tau) = E[X(t)X(t + \tau)]$	$R_X[m, k] = E[X_m, X_{m_k}]$

Theorem: $C_X(t, \tau) = R_X(t, \tau) - \mu_X(t)\mu_X(t + \tau)$
Stationary Process: A random process X(t0 is stationary if and only if for all sets of time constants $t, ..., t_m$ and any time difference τ, $f_{X(t_1),...,X(t_k)}(x_1, ..., x_k) = f_{X(t_1 + \tau),...,X(t_k + \tau)}(x_1, ..., x_k)$
Stationary Sequence: $f_{X_{n_1},...,X_{n_m}}(x_1, ..., x_m) = f_{X_{n_1 + k},...,X_{n_m + k}}(x_1, ..., x_m)$

Stationary Properties		Random Process	Random Sequence
	Expected Value	$\mu_X(t) = \mu_X$	$E[X_m] = \mu_X$
	Autocovariance	$C_X(t,\tau) = R_X(\tau) - \mu_X^2) = C_X(\tau)$	$C_X[m,k] = R_X(\tau) - \mu_X^2$
	Autocorrelation	$R_X(t,\tau) = R_X(0,\tau) = R_X(\tau)]$	$R_X[m,k] = R_X[0,k] = R_X[k]$

Wides Sense Stationary Properties		Random Process	Random Sequence
	Expected Value	$E[X(t)] = \mu_X$	$E[X_n] = \mu_X$
	Autocorrelation	$R_X(t,\tau) = R_X(0,\tau) = R_X(\tau)]$	$R_X[n,k] = R_X[0,k] = R_X[k]$

If $(X(t)/X_n$ is WSSP/WSRS	Processes	$R_X(0) \geq 0$ $R_X(\tau) = R_X(-\tau)$ $\lvert R_X(\tau)\rvert \leq R_X(0)$	Sequences	$R_X(0) \geq 0$ $R_X(k) = R_X(-k)$ $\lvert R_X(k)\rvert \leq R_X(0)$

Average Power: The average power of a WSSP $X(t)$ is $R_X(0) = E[X^2(t)]$

Random Signal Processing

LTI Filter Output Process: $X(t)$ is the input to a LTI filter and $Y(t)$ is the output. $Y(t)$ is the convolution of the sample function $X(t)$ with the impulse response $h(t)$

Theorem: If the input to LTI filter with impulse response is WSSP $X(t)$, then WSSP Output $Y(t)$ has the following

Mean Value	$\mu_Y = \mu_X H(0)$	Autocorrelation Function	$R_Y(\tau) = \int_{-\infty}^{\infty} h(u) \int_{-\infty}^{\infty} h(v) R_X(\tau + u - v) dvdu$

Power Spectral Density: For a WSSP $X(t)$, $R_X(\tau)$ and the power spectral density $S_X(f)$ are Fourier transform pairs

$S_X(f) = \mathcal{F}\{R_X(\tau)\}$	$R_X(\tau) = \mathcal{F}^{-1}\{S_X(f)\}$	$S_X(f) \geq 0$ for all f

Properties of $S_X(f)$	$E[X^2(t)] = \int_{-\infty} \infty S_X(f) df$	$S_X(f) = S_X(-f)$	$S_Y(f) = S_X(f)\lvert H(f)\rvert^2$ where H is the input response

Independent Processes: $X(t)$ $Y(t)$ are independent for any time sample of $t_1, .., t_n$, and $t'_1, ..., t'_m$,
$f_{X(t_1),...,X(t_n),Y(t'_1),...,Y(t'_m)}(x_1, ..., x_n, y_1, ..., y_m) = f_{X(t_1),...,X(t_n)}(x_1, ..., x_n,) f_{Y(t'_1),...,Y(t'_m)}(y_1, ..., y_m)$

Cross Correlation Function: $R_{XY}(t,\tau) = E[(X(t)Y(t+\tau)]$

Jointly WSSP: $R_{XY}(t, t+\tau) = R_{XY}(\tau)$ and $R_{XY}(\tau) = R_{XY}(-\tau)$

Cross Spectral Density: $S_{XY}(f) = \mathcal{F}\{R_{XY}(\tau)\}$

Input-Output Cross Correlation: When a WSSP $X(t)$ is the input to a LTI filter $h(t)$, the input-output cross correlation is $R_{XY}(t, t+\tau) = R_{XY}(\tau) = \int_{-\infty}^{\infty} h(u) R_X(\tau - u) du$. Both $X(t)$ and $Y(t)$ are jointly wide sense stationary.

Output Autocorrelation: When a wide stationary process $X(t)$ is the input to a LTI filter $h(t)$, the autocorrelation of the output $Y(t)$ is $R_Y(\tau) = \int_{-\infty}^{\infty} h(-w) R_{XY}(\tau - w) dw$.

Cross Spectral Properties: Let $X(t)$ be a WSS input to a LTI filter $H(f)$. The input $X(t)$ and output $Y(t)$ satisfy

$S_{XY}(f) = H(f) S_X(f)$	$S_Y(f) = H(f) S_{XY}(f)$

Gaussian Process: $X(t)$ is a Gaussian random process if the joint PDF of $X(t_1,), ..., X(t_k)$ has the multivariate Gaussian density $f_{X(t_1),...X(t_k)} = \frac{1}{(2\pi)^{k/2}\lvert C\rvert^{1/2}} \exp\{-\frac{1}{2}(X - \mu x)^\tau C^{-1}(X - \mu x)\}$

Theorem: If $X(t)$ is WSS Gaussian process, then $X(t)$ is a stationary Gaussian processs. Theorem: $X(t)$ is a Gaussian RV if $Y = \int_0^T g(t) X(t) dt$ is a Gaussian random variable for every $g(t)$ such that $E[Y^2] < \infty$.

Properties of Gaussian Process: Passing a stationary Gaussian process $X(t)$ through a linear filter $h(t)$ yields as the output Gaussian random process $Y(t)$ with the following properties,

Mean	$\mu_Y = \mu_X H(0)$	Autocorrelation	$R_Y(\tau) = \int_{-\infty}^{\infty} h(u) \int_{-\infty}^{\infty} h(v) R_X(\tau + u - v) dvdu$

White Gaussian Noise Processes

Noise is an unpredictable waveform that we model as stationary Gaussian random process $W(t)$. Noise has no DC component. $E[W(t_t 1)] = \mu_W = 0$ and $R_W(\tau) = 0$.

Power Spectral Density of $W(t)$ $S_W(f)$ is constant. The constant is 0 unless $R_W(\tau) = \frac{N_0}{2}\delta(\tau)$. N_0 is the the power per unit bandwidth of $W(t)$.

Average Noise Power: $E[W^2(t)] = R_W(0) = \infty$

Noise process output: $Y(t) = \int_0^t h(t - \tau) W(\tau) d\tau =$ Constant