$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx = \mu_x$, $\sum_i x_i \cdot P(x_i)$ = discrete

$E[Z] = E[X+Y] = E[X] + E[Y]$, If X and Y are S.I., $E[XY] = E[X] \cdot E[Y]$. Also, $E[X_1 X_2 \cdots] = E[X_1]E[X_2]\cdots$

$Var(X) = E[(X-\mu_x)^2]$ = central moment, 2nd $\rightarrow E[X^2 - 2X\mu_x + \mu_x^2] = E[X^2] - 2E[X]\cdot\mu_x + \mu_x^2$

$= E[X^2] - \mu_x^2 = E[X^2] - E[X]^2$

$E[X]$ = linear, $E[X] = E[X^+] - E[X^-]$ where $E[X^+] = \int_0^{\infty} x \cdot f_X(x)\,dx$, $E[X^-] = \int_0^{\infty} x \cdot f_X(-x)\,dx$

Lotus $= E[Y] = E[g(x)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\,dx$

$Var(c) = 0$, $Var(X+c) = Var(X)$, $Var(cX) = c^2 Var(X)$. If S.I., $Var[XY] = Var[X] + Var[Y]$

$Var[X+Y] = Var[X] + Var[Y] + 2Cov[X,Y]$

$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_Y)]$, $Var[X] = cov(X,X) \rightarrow Cov(X,Y) = E[XY] - \mu_x\mu_y$.

If $\mu_x = 0$ or $\mu_y = 0$, then $cov(X,Y) = E[XY]$. If X and Y = S.I., then $Cov(X,Y) = 0$

Random variable = orthogonal if $E[XY] = 0$.

inner product

ρ_{XY} = correlation coefficient. $= Cov(X,Y) / \sigma_X \cdot \sigma_Y$ where it ranges from -1 to 1. σ_x = std !!

Joint expected value $= E[Y \mid X = x] = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y|x)\,dy$ → notice this function will become a random variable too. And it is a random variable based on X. $E_X[E_Y[Y|X]] = E[Y]$!!!

Min L2 error has multiple types. ① Constant $= E[X]$. ② Affine function $= aY + b = \hat{X}$

$a = Cov(X,Y)/\sigma_Y^2 = \rho \frac{\sigma_X}{\sigma_Y}$, $b = \mu_X - a\mu_Y$ → $\hat{X} = \rho\sigma_X \cdot \left(\frac{y-\mu_y}{\sigma_y}\right) + \mu_X$

③ Use condition, $\hat{X} = E[X|Y]$ → use Y to estimate X

$Z = X + Y$ where X and Y = Gaussian then Z is gaussian too. We can use the formula above to obtain the new mean and variance.

$f_{WZ} = f_{XY} \cdot 1/|J_{WZ}|$ where $J_{WZ} = \begin{array}{c|cc} & X & Y \\ \hline W & \frac{\partial W}{\partial X} & \frac{\partial W}{\partial Y} \\ Z & \frac{\partial Z}{\partial X} & \frac{\partial Z}{\partial Y} \end{array}$

$f_Y(y) = f_X(x) \cdot 1/|dY/dX|$ where x = inverse of Y. If Y has multiple corresponding x then we sum all possible density.

If $Z = X + Y$ where X and Y = S.I. we can get it by using convolution integral $f_Z(z) = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x)\,dx$

characteristic function $\Phi_X(w) = E[e^{jwX}] = \int_{-\infty}^{\infty} f_X(x) \cdot e^{jwx} \cdot dx$ → fourier at $-w$.

Moment generating function $\theta_X(t) = E[e^{tX}]$ $\int_{-\infty}^{\infty} f_X(x) \cdot e^{tx} \cdot dx$ → laplace at $-t$

Notice $t = jw$ or $s = jw$ in this form. Notice if we take $\left(\frac{d}{dt}\right)^n$ or $\left(\frac{1}{j}\frac{d}{dw}\right)^n$ on those functions and set $w=0$ or $t=0$, we get $E[X^n]$. Notice we can also use this to avoid convolution

Markov Bound $= P(X \geq a) \leq E[X]/a$ ← this only for $X \geq 0$.

Chebyshev Inequality: $P(|X-\mu_x| \geq a) \leq \sigma_x^2/a^2$ ← need μ_x and σ_x

Chernoff $= P(Z \geq a) = \min_{s>0} e^{-as} \cdot \theta_Z(s)$, usually take $\frac{d}{ds}$ and set it to $= 0$.

MSE for linear if using Y to estimate X, $Var(X) - \frac{Cov(X,Y)^2}{Var(Y)}$

For X given Y, best: $E[Var(X|Y)]$