**Random variables**
$f_X$ – density of X
$F_X$ – distrubution function of X
$\mu_X$ – law of X

## Distribution function and density

1. Discrete:

$\mu_X[\mathcal{B}] = P[X \in \mathcal{B}]$

$p_X(x) = P[X = x] = \mu_X(\{x\})$

$X = (X_1, X_2, \dots, X_n)$ iid $\Leftrightarrow p_X(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$

2. Absolutely continuous:

$F_X(x) = P[X \le x] = \int_{-\infty}^{x} f_X(t)dt$ → int(0,x) if positive RV $\qquad f_X(x) = \frac{\partial F_X(x)}{\partial x}$

if $x < y \Rightarrow F_X(x) \le F_Y(y)$

$P[X = x] = F_X(x) - \lim_{y \uparrow x} F_X(y)$

$\int_{-\infty}^{+\infty} f_X(t)dt = 1$

$P[Z^n < z] = \begin{cases} & x < \alpha^2 \\ F_Z(\sqrt{z}) & \alpha^2 \le x < \beta^2 \\ & x \ge \beta^2 \end{cases}$

open or closed

$P[a < X < b] = P[X \in [a, b]] = F_X(b) - F_X(a) = \int_a^b f_X(t)dt$

$X = (X_1, X_2, \dots, X_n)$ iid $\Leftrightarrow f_X(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$

## Expectation

$E[X] = \begin{cases} \sum_{k \in R} k\,p_X(k) \ (discrete) \\ \int_{-\infty}^{+\infty} x f_X(x)dx \ (abs.cont) \end{cases}$

$E[g(x)] = \begin{cases} \sum_{k \in R} g(k)p_X(k) \ (discrete) \\ \int_{-\infty}^{+\infty} g(x)f_X(x)dx \ (abs.cont.) \end{cases}$

**Order statistics**

$X_{(1)} = \min\{X_i\}, X_{(3)} = \max\{X_i\}$

$F_{X_{(1)}}, F_{X_{(3)}} \rightarrow$ ☆

*If* $x_1 > 0 \quad x_3 > 0$:

$P[X_{(1)} > x_1, X_{(3)} \le x_3]$
$= P[x_1 < X_{(i)} \le x_3]$
$= \prod_{i=1}^{n} P[x_1 < X_i \le x_3]$
$= [F_{X(1)}(x_1) - F_{X(3)}(x_3)]^n$

$\boldsymbol{F}_{X(3)}(x_1, x_3) = P[X_{(3)} < x_3] - P[X_{(1)} > x_1, X_{(3)} < x_3]$
$= F_{X(3)}(x_3) - [F_{X(1)}(x_1) - F_{X(3)}(x_3)]^n$

Properties:

1. $x \le y \Rightarrow E[X] \le E[Y]$ (monotonicity)
2. $P[X = c] = 1 \Rightarrow E[X] = c$
3. $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y]$ (linearity)

$E\left[\sum_i X_i\right] = \sum_i E[X_i]$

$E[XY] = E[X] \cdot E[Y]$ if $X \perp\!\!\!\perp Y$

$E[X] = -\int_{-\infty}^{0} F_X(x)dx + \int_{0}^{+\infty} (1 - F_X(x))dx$

abs. cont. $\rightarrow E[X] = \int_0^{+\infty} P[X > x]dx = \int_0^{+\infty} (1 - F_X(x))dx$

**Positive RVs ($x \ge 0$)**

discrete $\rightarrow E[X] = \sum_{n=1}^{+\infty} P[X \ge n] = \sum_{n=1}^{+\infty} k \cdot P[X = k]$

## Variance

$Var[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2$

$Var[X + Y] = Var[X] + Var[X - Y]$

$Var[\alpha X] = \alpha^2 Var[X]$

**Covariance**

$Cov(X, Y) = E[XY] - E[X] \cdot E[Y]$
$Cov(X, X) = Var[X]$
$Cov(X, Y) = Cov(Y, X)$
$Cov(cX, Y) = c \cdot Cov(X, Y)$
$Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$
if $X \perp\!\!\!\perp Y \Rightarrow Cov[X, Y] = 0$ (uncorrelated)

**Conditional Expectation**

D: $p_{X|Y}(x|y) = P[X{=}x|Y{=}y] = \frac{P[X=x, Y=y]}{P[Y=y]}$

$E[X|Y{=}y] = \Sigma\, x\, p_{X|Y}(x|y)$

AC: $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

$E[X|Y{=}y] = \int x\, f_{X|Y}(x|y)\, 1_{(\dots,\dots)}$

Tower property: $E[\,E[X|Y{=}y]\,] = E[X]$

# DISCRETE

### 1. Bernoulli: X~Be(p) = Bin(1,p)

$x = \begin{cases} 0, & 1 - p = q \\ 1, & p \end{cases}$

$p_X(k) = P[X = k] = p^k(1 - p)^{1-k}$
$E[X] = E[X^2] = p$
$Var[X] = pq = p(1 - p)$
$m_X(t) = (1 - p) + pe^t$
$\varphi_X(u) = (1 - p) + pe^{iu}$

Total probability:

$\mathcal{P}[Y{=}k] =$
$\mathcal{P}[Y{=}k \mid X_0{=}0]\ \mathcal{P}[X_0{=}0]$
$+ \mathcal{P}[Y{=}k \mid X_1{=}1]\ \mathcal{P}[X_1{=}1]$

$m(t) = 1 - p + pe^t \le e^{p(e^t - 1)}$

since $1 + x \le e^x$

Upper: $P[\overline{X_n} \ge p(1 + \delta)] \le e^{-pn\frac{\delta^2}{2+\delta}} \rightarrow$ use inequality (4) and (1) $\qquad \bar{t}_U = log(1 + \delta)$

Lower: $P[\overline{X_n} \le p(1 - \delta)] \le e^{-pn\frac{\delta^2}{2}} \rightarrow$ use inequality (10) $\qquad \bar{t}_L = log(\frac{1}{1-\delta})$

### 2. Binomial: X~Bin(n, p)

*# successes in n trials* $\qquad E[X^2] = np(1\text{-}p + np)$ $\qquad \sum_{k=0}^{n}$

$p_X(k) = P[X = k] = \binom{n}{k} p^k(1 - p)^{n-k}$ $\qquad$ **X+Y ~ Bin(2n, p)**

$E[X] = np$ $\quad \rightarrow$ *easy to prove using* $n \cdot Bin(1,p)$

$Var[X] = np(1 - p)$ ↗

$m_X(t) = [(1 - p) + pe^t]^n$ $\qquad$ Upper: $P[\overline{X_n} \ge p(1 + \delta)] \le e^{-\frac{\delta^2 p}{3}}$

$\varphi_X(u) = [(1 - p) + pe^{iu}]^n$ $\qquad$ Lower: $P[\overline{X_n} \ge p(1 - \delta)] \le e^{-\frac{\delta^2 p}{2}}$

### 3. Geometric: X~Geo(p)

*first success on k-th trial* $\qquad \sum_{k=1}^{+\infty}$

$p_X(k) = P[X = k] = p(1 - p)^{k-1}, k \ge 1, p \in (0,1]$

$P[X \le k] = 1 - (1 - p)^k$

$E[X] = \frac{1}{p}$ $\qquad E[X^2] = (2\text{-}p)\, p^{-2}$

$Var[X] = \frac{1 - p}{p^2}$

$m_X(t) = \frac{p}{e^{-t} - 1 + p}$, $\quad t < -\log(1 - p)$

$\varphi_X(u) = \frac{p}{e^{-iu} - 1 + p}$

### 4. Hypergeometric: X~Hp(n, N, m)

*n draws (without replacement) from an urn with m red, N−m blue*
*X: "number of red balls among the n drawn"*

With replacement: $X \sim Bin(n, \frac{m}{N})$ $\qquad Var[X] = \frac{nm(N-m)}{N^2}\frac{N-n}{N-1}$

$p_X(k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$

$\mathbb{P}[\mathbb{S}_{n,p} \geqslant n(p + t)] \leqslant \exp(-nD_{KL}(p + t, p)) \leqslant \exp(-2nt^2)$

$\mathbb{P}[\mathbb{S}_{n,p} \geqslant n(p + t)] \leqslant \exp(-nD_{KL}(p + t, p)) \leqslant \exp\left(-n\frac{t^2}{2p(1-p)}\right)$

when $p \geqslant 1/2$

$E[X] = n\frac{m}{N}$

### 5. Poisson: X~Pois(λ)

$p_X(k) = P[X = k] = e^{-\lambda}\frac{\lambda^k}{k!}$ $\qquad E[X^2] = \lambda\,(\lambda{+}1)$

$E[X] = Var[X] = \lambda$

$X \sim Bin(n, p)$ $\quad$ *n large, p small* $\Rightarrow$ *approximate with Pois with* $\lambda = np$

$m_X(t) = e^{\lambda(e^t - 1)}$ $\qquad \bar{t}_U = log(1 + \varepsilon)$

$\varphi_X(u) = e^{\lambda(e^{iu} - 1)}$ $\qquad \bar{t}_L = -log(1 - \varepsilon)$

if $X \sim Pois(\lambda) \perp\!\!\!\perp Y \sim Pois(\mu) \Rightarrow X + Y \sim Pois(\lambda + \mu)$

Upper: $P[\overline{X_n} \ge \lambda(1 + \varepsilon)] \le e^{-n\lambda\frac{\varepsilon^2}{2+\varepsilon}} \rightarrow$ use inequality (11)

Lower: $P[\overline{X_n} \le \lambda(1 - \varepsilon)] \le e^{-n\lambda\frac{\varepsilon^2}{2}} \rightarrow$ use inequality (10)

# ABSOLUTELY CONTINUOUS

### 6. Uniform: X~U(a, b)

$f_X(x) = \frac{1}{b - a}\mathbb{1}_{(a,b)}(x)$ $\qquad E[X^2] = \frac{(b\text{-}a)^2 + 3(b\text{-}a)^2}{12}$

$F_X(x) = \begin{cases} 0, & x < a \\ \frac{x - a}{b - a}, & a \le x < b \\ 1, & x \ge b \end{cases}$

$E[X] = \frac{b + a}{2}$

$Var[X] = \frac{(b - a)^2}{12}$

$m_X(t) = \begin{cases} \frac{e^{tb} - e^{ta}}{t(b - a)}, & t \neq 0 \\ 1, & t = 0 \end{cases}$

$\varphi_X(u) = \begin{cases} \frac{e^{iub} - e^{iua}}{iu(b - a)}, & u \neq 0 \\ 1, & u = 0 \end{cases}$

☆

$Z = max\{X_1, \dots, X_n\}$
$\Longrightarrow F_Z(z) = P[X_1 < z] \cdot \dots \cdot P[X_n < z]$
$= F_Z(X_1) \cdot \dots \cdot F_Z(X_n)$
^ If all $X_i$ are the same distribution, then $F_Z(z) = [F_Z(X_i)]^n$

$Z = min\{X_1, \dots, X_n\}$
$\Longrightarrow F_Z(z) = 1 - (P[X_1 > z] \cdot \dots \cdot P[X_n > z])$
$= 1 - [(1 - P[X_1 < z]) \cdot \dots \cdot (1 - P[X_n < z])]$
$= 1 - [(1 - F_Z(X_1)) \cdot \dots \cdot (1 - F_Z(X_n))]$
^ If all $X_i$ are the same distribution, then $F_Z(z) = 1 - [1 - F_Z(X_i)]^n$
**! Careful with the min{ } of positive RVs in which you use 1 - CDF**

For U(-1,1):

Upper: $P[\overline{X_n} \ge \varepsilon] \le e^{-n\frac{\varepsilon^2}{2}} \rightarrow$ use $\frac{\sinh\varepsilon}{\varepsilon} \le 1 + \frac{\varepsilon^2}{2}$ *and* inequalitiey (4) and/or (5)

Lower: $P[\overline{X_n} \le -\varepsilon] \le e^{-n\frac{\varepsilon^2}{2}}$ ↗

### 7. Exponential: X~Exp(λ), λ > 0

$f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{(0,+\infty)}(x)$ $\qquad \mathcal{E}[X^2] = \frac{2}{\lambda^2}$

$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \ge 0 \end{cases}$ $\qquad \mathcal{E}[X^3] = \frac{6}{\lambda^3}$

$E[X] = \int_0^{+\infty} P[X \ge x]dx = \int_0^{+\infty} e^{-\lambda x}dx = \frac{1}{\lambda}$ $\qquad \mathcal{E}[X^4] = \frac{24}{\lambda^4}$

$Var[X] = \frac{1}{\lambda^2}$

Memoryless property: $P[T > s + t | T > s] = P[T > t]$, $T \sim Exp(\lambda)$

$m_X(t) = \frac{\lambda}{\lambda - t}$, $\quad t < \lambda$ (not defined otherwise)

$\varphi_X(u) = \frac{\lambda}{\lambda - iu}$ $\qquad \bar{t}_U = \frac{\varepsilon\lambda^2}{1 + \varepsilon\lambda}$ $\qquad \bar{t}_L = \frac{\varepsilon\lambda^2}{1 - \varepsilon\lambda}$

Upper tail: $P[\overline{X_n} \ge \lambda(1 + \delta)] \le e^{-n\frac{\delta^2}{2(1+\delta)}}$ for $\delta > 0$

Lower tail: $P[\overline{X_n} \le \lambda(1 - \delta)] \le e^{-n\frac{\delta^2}{2}}$ for $0 < \delta < 1$

### 8. Normal (Gaussian): X~N(μ, σ²)

$E[Z^2] = \mu^2 + \sigma^2$

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

$E[Z] = \mu$

$Var[Z] = \sigma^2$

$m_X(t) = e^{t\mu + \frac{t^2\sigma^2}{2}}$

$\varphi_X(u) = e^{iu\mu - \frac{u^2\sigma^2}{2}}$

Standard normal: $Z = \frac{x-\mu}{\sigma} \sim N(0,1)$

$F_Z(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$, $\Phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

$F_X(x) = P[X \le x] = P\left[Z \le \frac{x - \mu}{\sigma}\right]$

$a + bX \sim N(a + b\mu, b^2\sigma^2)$

if $X \sim N(\mu_1, \sigma_1^2) \perp\!\!\!\perp Y \sim N(\mu_2, \sigma_2^2) \Rightarrow X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Upper tail: $P[\overline{X_n} \ge \varepsilon] \le e^{-n\frac{\varepsilon^2}{2}} \rightarrow$ use $P[X > x] < \frac{1}{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ and $P[X > x] = \int_x^{+\infty}\Phi(t)dt$

### Chi-squared: $X \sim Z^2(k)$

$E[X]{=}k \qquad Var[X]{=}2k \qquad m_X(t) = (1 - 2t)^{-\frac{k}{2}} = \sqrt{\frac{1}{(1-2t)^k}}$ $\qquad k = 1,\ t < 1/2 \qquad \bar{t} = \frac{1}{2}\frac{\varepsilon}{\varepsilon + 1}$

$E[X^2] = k\,(k{+}2)$

Careful! When you compute CDF in a max\min exercise, the CDF can become nonzero for the outliers, hence the new E[Z] will be the integral from -inf to $\alpha$ (or 0 to $\alpha$ for positive RVs) of the left outlier plus the integral from $\alpha$ to $\beta$ of z.f (z) plus the integral from $\beta$ to +inf of the right outlier.

The integrals of the outliers can converge hence change your E[Z], so always compute.
The same process must be repeated if we are asked to compute exponential variations of the CDF for computing Var[Z].
Example: exam 24/01/2023.

## Moment generating function

$$m_X(t) = E[e^{tX}] = \begin{cases} \sum_x e^{tx} p_X(x) \ (discrete) \\ \int_{-\infty}^{+\infty} e^{tx} f_X(x)dx \ (abs. cont.) \end{cases}$$

$$\left.\frac{dm_X(t)}{dt}\right|_{t=0} = m'(t=0) = E[X]$$

$E[X^2] = m''(t=0)$ $\quad E[X^3] = m'''(t=0)$

## Characteristic function

$\varphi_X(u) = E[e^{iuX}] = E[\cos(uX)] + iE[\sin(ux)]$

$$\left.\frac{d^n\varphi_X(u)}{du^n}\right|_{u=0} = i^n E[X^n]$$

$\varphi_X = \varphi_Y \Leftrightarrow \mu_X = \mu_Y$

if $X \perp\!\!\!\perp Y \Rightarrow$ $\varphi_{(X,Y)}(u,v) = \varphi_X(u)\varphi_Y(v)$; $m_{(X,Y)}(s,t) = m_X(s)m_Y(t)$

$\varphi_{aX+bY}(u) = \varphi_X(au)\varphi_Y(bu)$; $m_{aX+bY}(t) = m_X(at)m_Y(bt)$

*for a given* $f_X(x)$ *we first calculate* $\varphi_X(u)$*, and* $m_X(t) = \varphi_X(-it)$

## Bounds

Markov: $P[X \geq \varepsilon] \leq \frac{E[X]}{\varepsilon}$

Chebischev: $P[|X-\mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2}$

### Chernoff bounds

$P[\overline{X}_n \geq \mu + \varepsilon] \leq e^{-g(t^*)n} \rightarrow$ exp. decay for the upper tail

$g(t) = t(\mu+\varepsilon) - \log(m(t))$

$P[\overline{X}_n \leq \mu - \varepsilon] \leq e^{-h(t^*)n} \rightarrow$ exp. decay for the lower tail

$h(t) = -t(\mu-\varepsilon) - \log(m(-t))$

CLT:

$$Z_n = \sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow[n\to\infty]{D} Z \sim N(0,1)$$

Hoeffding inequality (upper bound):

$S_n = \sum_{i=1}^n Z_i, Z_i \epsilon [a,b]$ $\quad P[S_n - E[S_n] \geq t] \leq e^{-\frac{2t^2}{n(b-a)^2}}$

## Basic probability

$(A \cup_\cap B)^c = A^c \cap_\cup B^c$

$P[B] = P[A \cap B] + P[A^c \cap B]$

$P[A \cup B] = P[A] + P[B] + P[A \cap B]$

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

$P[A \cap B] = P[A]P[B] \Leftrightarrow A \perp\!\!\!\perp B$

$\bigcap_{i=1}^n A_i = \prod_{i=1}^n P[A_i]$

| | Order Doesn't Matter | Order Matters |
|---|---|---|
| With Replacement | $\frac{(n+k-1)!}{k!\,(n-1)!}$ | $n^k$ |
| Without Replacement | Combination $\frac{n!}{(n-k)!\,k!}$ | Permutation $\frac{n!}{(n-k)!}$ |

## Jointly distributed

$F_{X,Y}(x,y) = P[X \leq x, Y \leq y]$

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

$X \perp\!\!\!\perp Y \Leftrightarrow$ $P[X\epsilon A, Y\epsilon B] = P[X\epsilon A]P[Y\epsilon B]$; $F_{X,Y}(x,y) = F_X(x)F_Y(y)$

$F_X(a) = P[X \leq a, y < +\infty]$

Probability of getting from k to anywhere in n steps?

$$\mathcal{E}[X_n | X_0 = k] = \sum_{i=1}^{\infty} i \cdot \mathcal{P}[X_n = i | X_0 = k]$$

## Markov Chains

P is a transition matrix if $\sum_x P_{yx} = 1, \forall y \in S$ (sum of the elements of each row=1)

P is a stochastic matrix if $\sum_x P_{yx} = 1, \forall y \in S$ and $P_{yx} \geq 0\ \forall x, y \in S$

Chapman-Kolmogorov:

$P_{yx}^{m+n} = \sum_{z \in S} P_{yz}^m P_{zx}^n$ - *The path from state y to x in (m+n) steps can be broken down in two paths: from y to z in m steps, and from z to x in n steps*

STATE CLASSIFICATION

1. Communication:

Irreducible iff all states belong to one communication class. $S = C_1$

Reducible iff there are two or more com. classes. $S = \bigcup_{i=1}^n Ci$

2. Periodicity:

$d(k) = gcd\{n \geq 1 : P_{ii}^n > 0\} = 1 \Rightarrow$ aperiodic

*If at least one of the states is aperiodic, then the remaining states are aperiodic.* $\Rightarrow$ *MC is aperiodic*

e.g.: $P_{11}^2 > 0, P_{11}^3 > 0 \rightarrow$ state 1 is aperiodic $\Rightarrow$ MC is aperiodic

3. Recurrence and transiency – *how often we visit the states*

$f_i = P[X_n = i \text{ for some } n \geq 1 | X_0 = i]$

$N_i = |\{n \geq 0 : X_n = i\}|$

Recurrent: $f_i = 1$ and $P[N_i = +\infty | X_0 = i] = 1$

Transient: $f_i < 1$ and $P[N_i < +\infty | X_0 = i] = 0$

A class C is closed $\Leftrightarrow$ all its states are recurrent

$T_i = \inf\{n \geq 1 : X_n = i\}$ – *time of first visit to state i*

$m_i = E[T^i | X_0 = i]$

Positive recurrent if $m_i < +\infty$ (*from i it takes finite number of steps to go back to i*)

Null recurrent if it's recurrent and $m_i = +\infty$.

Aperiodic, positive recurrent state is called ergodic.

If all states are ergodic, the MC is ergodic.

Ergodic theorem:

For an irreducible and ergodic MC, exists $(\pi_j)_{j \in S}$ which is a distribution ($\pi_j \geq 0, \sum_{j \in S} \pi_j = 1$), called the limit distribution.

$$(a) \begin{cases} \pi = \pi P \\ \sum_{j \in S} \pi_j = 1 \end{cases}$$

(b) $\pi_j = \frac{1}{m_j}$

$m_j$ – *the expected time to go back to j, after starting from j*

$\pi_j$ – *long run proportion of time the chain spends in state j*

(c) $\pi_j = \lim_{n \to +\infty} \frac{\#visits\ to\ j\ by\ time\ n}{n}$

$\pi$ is unique invariant distribution if MC is ergodic and $|S| < +\infty$

$\pi$ is reversible if:

(DB) $\pi_i P_{ij} = \pi_j P_{ji}, \forall i, j \in S, i \neq j$

If $\pi$ is reversible $\Rightarrow \pi$ is invariant.

## Random Walks:

$v_i$ = number of direct connections to other states.

$$\pi = \frac{v_1, v_2, v_3, \ldots}{\sum v_i} = [\pi_1, \pi_2, \pi_3, \ldots]$$

- If the graph is connected, the RW is irreducible and reversible.
- If $v_{i,i} > 0$ in both 2 and 3 steps, the RW is aperiodic.

## Extras:

$P[X < Y] = \int_0^{+\infty} f_Y(y)\left(\int_0^y f_X(x)dx\right)dy = \int_0^{+\infty} f_X(x)\left(\int_x^{+\infty} f_Y(y)dy\right)dx$ $\quad$ |X,Y abs. cont.

$E[X Y^2] = E[X]\, E[Y^2]$

$E[X(Y+1)] = E[X]\,(E[Y]+1)$

$E[X|X < a] = \int_0^{+\infty} P[X > x | X < a]dx = \int_0^a \frac{P[x < X < a]}{P[X < a]}dx$ $\quad$ $P[X<a]-P[X<x]$ $\quad$ $x<a$ for $e^{(0,a)}$

$P[X > Y] = P[Y < X] = \sum_{k=1}^n P[X > Y | Y = k] \cdot P[Y = k] = \sum_{k=1}^n P[X > k] \cdot P[Y = k] = \sum_{k=1}^n [1 - F_{X(k)}]\, p_{Y(k)}$ $\quad$ |Y discrete

$E[X^{Y+1}] = \sum_{k=1}^n E[X^{Y+1} | Y = k]P[Y = k] = \sum_{k=1}^n E[X^{k+1}]P[Y = k]$

## Sums:

$\sum_{i=0}^N a^i = \frac{1-a^{N+1}}{1-a}$ $\quad$ $\sum_{i=1}^N a^i = a\frac{a^N - 1}{a-1}$ $\quad$ $\sum_{k=1}^n a^{k-1} = \sum_{t=0}^n a^t$ $\quad t = k-1$

$\sum_{k=0}^{+\infty} aq^k = \frac{a}{1-q}, |q| < 1$ $\quad$ $\sum_{k=0}^{\infty} \frac{1}{k+1}\frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-\lambda}}{\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{k+1}}{(k+1)!} = \frac{e^{-\lambda}}{\lambda}(e^{\lambda} - 1)$

$\sum_{n=0}^{+\infty} \frac{f^n(a)}{n!}(x-a)^n = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots$ (general Taylor expansion)

$\sum_{k=0}^{+\infty} \frac{x^k}{k!} = e^x \rightarrow \sum_{k=0}^n \frac{1}{k!} = e$

$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$ $\quad$ $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$

$e^{ix} = \cos(x) + i\sin(x)$

Useful Taylor expansions:

$\log(1+\varepsilon) = \varepsilon - \frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3} - \cdots$ $\quad$ $\log(1-\varepsilon) = -\varepsilon - \frac{\varepsilon^2}{2} - \frac{\varepsilon^3}{3} - \cdots$

Derivatives and integrals:

$f'(g(x)) = f'(g(x))g'(x)$ $\quad$ (Chain rule)

$\int udv = uv - \int vdu$ $\quad$ (Integration by parts)

$\frac{d}{dx}x = 1; \int 1dx = x$

$\frac{d}{dx}x^n = nx^{n-1}; \int x^n dx = \frac{x^{n+1}}{n+1}$ $\quad$ $\int_a^b e^{-\sqrt{x}}\,dx = \int_{\sqrt{a}}^{\sqrt{b}} 2t \cdot e^{-t}\,dt$

$\frac{d}{dx}e^{kx} = ke^{kx}; \int e^{kx}dx = \frac{1}{k}e^{kx}$

$\frac{d}{dx}\ln(x) = \frac{1}{x}; \int \frac{1}{x}dx = \ln(x); \int \ln(x)\,dx = x\ln(x) - x$

$\frac{d}{dx}a^x = a^x\ln(a); \int a^x dx = \frac{a^x}{\ln(a)}$

$\frac{d}{dx}\sin(x) = \cos(x); \int \cos(x)dx = \sin(x)$

$\frac{d}{dx}\cos(x) = -\sin(x); \int \sin(x)dx = -\cos(x)$

$\frac{d}{dx}\tan(x) = \frac{1}{\cos^2(x)}; \int \sin(x)dx = -\cos(x)$

$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$ $\quad$ $\int_{-\infty}^{\infty} e^{-a(x+b)^2}\,dx = \sqrt{\frac{\pi}{a}}$

To prove an inequality:

*if f(0)=g(0) and f'(0)=g'(0) and* ***f''(x) ≥ g''(x)****, then* ***f(x)≥g(x)***

… simple inequalities for Chernoff Bounds

(1) $\log(1+\varepsilon) \geq \frac{2\varepsilon}{2+\varepsilon}$ $\quad$ (4) $-\log(1+\varepsilon) \geq -\varepsilon$ $\quad$ (5) $\Leftrightarrow 1+\varepsilon \leq e^{\varepsilon}$ $\quad$ (6) $-\log(1-\varepsilon) \geq \varepsilon + \frac{\varepsilon^2}{2}$ $\quad$ (10) $(1-\varepsilon)\log(1-\varepsilon) + \varepsilon \geq \frac{\varepsilon^2}{2}$

(7) $-\log(1+\varepsilon) \geq \frac{\varepsilon^2}{2(1+\varepsilon)} - \varepsilon$ $\quad$ (11) $(1+\varepsilon)\log(1+\varepsilon) - \varepsilon \geq \frac{\varepsilon^2}{2+\varepsilon}$

(2) $\log(1+\varepsilon) \geq \varepsilon - \frac{\varepsilon^2}{2}$ $\quad$ (3.1) $\log(1+\varepsilon) \leq \varepsilon$ $\quad$ (3.2) $\Leftrightarrow e^{\varepsilon} \geq 1+\varepsilon$ $\quad$ (8) $-\log(1-\varepsilon) \geq \varepsilon - \frac{\varepsilon^2}{2}$ $\quad$ (9) $-\log(1+\varepsilon) \geq \frac{\varepsilon^2}{2(1+\varepsilon)} - \varepsilon$

## Proving infiniteness of integrals (trick)

$$\int_{-\infty}^{+\infty} f(x)\,g(x)\,dx \geq \int_a^b f(x)\,g(x)\,dx \geq \int_a^b f(x)\,\underbrace{\min_{a,b} g(x)}\,dx = \min_{a,b} g(x) \int_a^b f(x)\,dx$$

Pick some easy a and b interval depending on your functions | pick either *a* or *b* that minimizes g(x) and plug it in | f(x)*g(x) is clearly ≥ f(x)*min g(x) | *min g(x)* becomes a constant that we can place behind the integral | check if the integral of f(x) is infinite |

If it is infinite, then our initial integral is infinite (diverges) too since it is bigger or equal than our infinite final integral.