PMF

Random: $\mathcal{R}_n$

- $p_i = \dfrac{1}{n}$ for Ω = $\{1,2,\ldots,n\}$.
- Ex
  - classical game of chance / classical probability drawing at random
  - fair gaming devices (well-balanced coins and dice, well shuffled decks of cards)
  - high-rate coded digital data
  - experiment where
    - there are only n possible outcomes and they are all equally probable
    - there is a balance of information about outcomes

Bernoulli: $\mathcal{B}(1,p)$

- Ω = {0,1}; $p_0 = q = 1\text{-}p$, $p_1 = p$
- $EX = 1p + 0(1-p) = p$
- $EX = 1p + 0(1-p) = p$

  $E\left[X^2\right] = p1^2 + (1-p)0^2 = p$

  $\operatorname{Var}(X) = E\left[X^2\right] - (EX)^2 = p - p^2 = p(1-p)$

  Alternatively, $\operatorname{Var}(X) = p(1-p)^2 + (1-p)(0-p)^2 = p(1-p)(1-p+p) = p(1-p)$.

Binomial: $\mathcal{B}(n,p)$

- $p_i = \binom{n}{i} p^i (1-p)^{n-i}$ for $\Omega = \{0,1,2,\ldots,n\}$
- X is the number of success in n Bernoulli trials and hence the sun of n independent, identically distributed Bernoulli r.v.
- $\Phi_X(u) = \left(1-p+pe^{ju}\right)^n$

  Pf. $\Phi_X(u) = E\left[e^{juX}\right] = \sum_{i=0}^{n} e^{jui}\binom{n}{i} p^i (1-p)^{n-i} = (1-p)^n \sum_{i=0}^{n}\binom{n}{i}\left(\dfrac{pe^{ju}}{1-p}\right)^i$

  $= (1-p)^n\left(1+\dfrac{pe^{ju}}{1-p}\right)^n = \left(1-p+pe^{ju}\right)^n$

- $EX = np$

Pf. Method 1 $EX=\sum_{i=0}^{n} i\binom{n}{i}p^i(1-p)^{n-i}=(1-p)^n\sum_{i=0}^{n} i\binom{n}{i}\left(\frac{p}{1-p}\right)^i$

$$=(1-p)^n n\frac{p}{1-p}\left(\frac{p}{1-p}+1\right)^{n-1}=np$$

Method 2 $\frac{d}{du}\Phi_X(u)=n\left(1-p+pe^{ju}\right)^{n-1}jpe^{ju}$. $EX=-j\frac{d}{du}\Phi_X(u)\Big|_{u=0}=np$.

- $EX^2=(np)^2+np(1-p)$

Pf. Method 1 $\sum_{k=0}^{n} k^2\binom{n}{k}r^k=nr(r+1)^{n-1}\left(1+(n-1)\frac{r}{r+1}\right)$

$$EX^2=\sum_{i=0}^{n} i^2\binom{n}{i}p^i(1-p)^{n-i}=(1-p)^n\sum_{i=0}^{n} i^2\binom{n}{i}\left(\frac{p}{1-p}\right)^i$$

$$=(1-p)^n n\frac{p}{1-p}\left(\frac{p}{1-p}+1\right)^{n-1}\left(1+(n-1)\frac{\frac{p}{1-p}}{\frac{p}{1-p}+1}\right)$$

$$=np\left(1+(n-1)p\right)=(np)^2+np(1-p)$$

Method 2 $\frac{d^2}{du^2}\Phi_X(u)=jnp\left(j\left(1-p+pe^{ju}\right)^{n-1}e^{ju}+(n-1)\left(1-p+pe^{ju}\right)^{n-2}e^{ju}pje^{ju}\right)$

$$=-np\left(\left(1-p+pe^{ju}\right)^{n-1}+p(n-1)\left(1-p+pe^{ju}\right)^{n-2}e^{ju}\right)e^{ju}$$

$$EX^2=-\frac{d^2}{du^2}\Phi_X(u)\Big|_{u=0}=np\left(1+p(n-1)\right)=np(1+np-p)$$

$$=(np)^2+np(1-p)$$

- $Var[X]=EX^2-(EX)^2=np(1-p)$
- $0\le p\le 1 \Rightarrow$ probability of single occurrence of ~ $\mathcal{B}(1,p)$
- If have $\mathcal{E}_1, \ldots, \mathcal{E}_n$ n unlinked repetition of $\mathcal{E}$ and event A for $\boldsymbol{\mathcal{E}}$, $\mathcal{B}(n,p)$ = the probability that A occurs k times in $\mathcal{E}_1, \ldots, \mathcal{E}_n$
- Maximum probability value happens at $k_{max}=\lfloor (n+1)p\rfloor \approx np$
  - When $(n+1)p$ is an integer, then the maximum is achieved at k_{max} and $k_{max}-1$.
- $\mathcal{B}(1,\frac{1}{2})\Rightarrow$ binomial that also random
- Ex
  - #heads in n toss of a coin (p = 0.5)

- #errors in n symbols of text
  (p = the probability of an error in a single symbol of text)

- Gaussian Approximation for Binomial Probabilities
  Binomial random variable becomes difficult to compute directly for large n because of the need to calculate factorial terms.

  $$\Pr[X=k] \simeq \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(k-np)^2}{2np(1-p)}}$$

  Proof Binomial random variable X is a sum of iid Bernoulli random variables (which have finite mean p and variance $p(1-p)$), by the Central limit theorem, its cdf approaches that of a Gaussian random variable $Y \sim \mathcal{N}(n\mu, n\sigma^2) = \mathcal{N}(np, np(1-p))$.

  $$\Pr[X=k] \simeq \Pr\left[k-\frac{1}{2} < Y < k+\frac{1}{2}\right] = \frac{1}{\sqrt{2\pi np(1-p)}} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} e^{-\frac{1}{2np(1-p)}(x-np)^2} dx$$

  $$\simeq \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{1}{2np(1-p)}(k-np)^2} 1$$

  , where the second approximation comes from appromiating the integral by the product of the integrand at the center of the interval (at $x = k$) of integration and the length of the interval of integration (1).

Geometric: $\mathcal{G}(\beta)$

- $p_i = (1-\beta)\beta^i$, $\Omega = \mathbb{N} \cup \{0\}$, $0 \le \beta < 1$
- $\beta = \frac{m}{m+1}$ where m = average waiting time/ lifetime
- $\Pr[X=k] = \Pr[k \text{ failures followed by a success}]$

  $= (\Pr[\text{failures}])^k \Pr[\text{success}]$
- $\Pr[X \ge k] = \beta^k$ = the probability of having at least k initial failure = the probability of having to perform at least k+1 trials.
- $\Pr[X > k] = \sum_{i=k+1}^{\infty} (1-\beta)\beta^i = (1-\beta)\frac{\beta^{k+1}}{1-\beta} = \beta^{k+1}$ = the probability of having at least k+1 initial failure.
- Memoryless property:
  - $\Pr[X \ge k+c \mid X \ge k] = \Pr[X \ge c]$ k, c > 0.
  - $\Pr[X > k+c \mid X \ge k] = \Pr[X > c]$ k, c > 0.

Pf. $\Pr\left[X \ge k + c \middle| X \ge k\right] = \dfrac{\Pr\left[X \ge k + c \wedge X \ge k\right]}{\Pr\left[X \ge k\right]} = \dfrac{\Pr\left[X \ge k + c\right]}{\Pr\left[X \ge k\right]} = \dfrac{\beta^{k+c}}{\beta^{k}} = \beta^{c}$.

- If a success has not occurred in the first k trials (already fails for k times), then the probability of having to perform at least j more trials is the same the probability of initially having to perform at least j trials.
- Each time a failure occurs, the system "forgets" and begins anew as if it were performing the first trial.

- Geometric r.v. is the only discrete r.v. that satisfies the memoryless property.
- Ex.
  - lifetimes of components, measured in discrete time units, when the fail catastrophically (without degradation due to aging)
  - waiting times
    - for next customer in a queue
    - between radioactive disintegrations
    - between photon emission
  - number of repeated, unlinked random experiments that must be performed prior to the first occurrence of a given event A
    - number of coin tosses prior to the first appearance of a 'head'

      number of trials required to observe the first success

Poisson: $\mathcal{P}(\lambda\tau)$

- $p_{\mathrm{i}} = e^{-(\lambda\tau)}\dfrac{(\lambda\tau)^{i}}{i!}$; $\Omega = \mathbb{N} \cup \{0\}$, $0 \le \lambda\tau = \alpha$
- $\Phi_X(u) = e^{\lambda\tau\left(e^{iu}-1\right)}$

$$\Phi_X(u) = Ee^{iuX} = \sum_{k=0}^{\infty} e^{-\lambda\tau}\frac{(\lambda\tau)^{k}}{k!}e^{iuk} = e^{-\lambda\tau}\sum_{k=0}^{\infty}\frac{\left(\lambda\tau e^{iu}\right)^{k}}{k!}$$

$$= e^{-\lambda\tau}e^{\lambda\tau e^{iu}} = e^{\lambda\tau\left(e^{iu}-1\right)}$$

- $EX = \lambda\tau$

  Pf. Method 1

$$EX = \sum_{i=0}^{\infty} e^{-\lambda\tau}\frac{(\lambda\tau)^{i}}{i!}i = \sum_{i=1}^{\infty} e^{-\lambda\tau}\frac{(\lambda\tau)^{i}}{i!}i + 0 = e^{-\lambda\tau}(\lambda\tau)\sum_{i=1}^{\infty}\frac{(\lambda\tau)^{i-1}}{(i-1)!}$$

$$= e^{-\lambda\tau}\lambda\tau\sum_{k=0}^{\infty}\frac{(\lambda\tau)^{k}}{k!} = e^{-\lambda\tau}\lambda\tau e^{\lambda\tau} = \lambda\tau$$

  Method 2

$$\frac{d}{du}\Phi_X(u) = i\lambda\tau e^{iu} e^{\lambda\tau\left(e^{iu}-1\right)}$$

$$EX = -i\frac{d}{du}\Phi_X(u)\bigg|_{u=0} = \lambda\tau$$

- If λ is the rate, τ denotes a certain time period or certain region in space, then $\alpha = EX$ is the average number of event occurrences in a specified time interval or region in space.
- $VAR(X) = \lambda\tau$

  Pf. Method 1

  Note that $\sum_{k=1}^{\infty} k\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{\lambda} + e^{\lambda}$.

  $$EX^2 = \sum_{i=0}^{\infty} e^{-(\lambda\tau)}\frac{(\lambda\tau)^i}{i!}i^2 = e^{-(\lambda\tau)}(\lambda\tau)\sum_{i=0}^{\infty} i\frac{(\lambda\tau)^{i-1}}{(i-1)!} = e^{-\lambda}\lambda\tau\left((\lambda\tau)e^{(\lambda\tau)} + e^{(\lambda\tau)}\right)$$

  $$= (\lambda\tau)^2 + \lambda\tau$$

  Method 2

  $$\frac{d^2}{du^2}\Phi_X(u) = -\lambda\tau e^{iu} e^{\lambda\tau\left(e^{iu}-1\right)} - (\lambda\tau)^2 e^{i2u} e^{\lambda\tau\left(e^{iu}-1\right)}$$

  $$EX^2 = -\frac{d^2}{du^2}\Phi_X(u)\bigg|_{u=0} = \lambda\tau + (\lambda\tau)^2$$

  $$VAR(X) = \left((\lambda\tau)^2 + \lambda\tau\right) - (\lambda\tau)^2 = \lambda\tau$$

- $\lambda\tau$ = mean/average #counts
  - τ = observation time
  - λ = an event intensity / rate of occurrence/current
- i.i.d. $N_k \sim \mathcal{P}(\lambda_k) \rightarrow N = \sum N_k \sim \mathcal{P}(\sum\lambda_k)$

  $$\Phi_N(u) = \prod_k \Phi_{N_k}(u) = \prod_k e^{\lambda_k\left(e^{iu}-1\right)} = e^{\left(\sum\lambda_k\right)\left(e^{iu}-1\right)}$$

- Most probable value $(i_{\max})$

	Most probable value $(i_{\max})$	Associated max probability
$0 < \alpha < 1$	0	$e^{-\alpha}$
$\alpha \in \mathbb{N}$	$\alpha - 1, \alpha$	$\frac{\alpha^{\alpha}}{\alpha!}e^{-\alpha}$
$\alpha \geq 1,\ \alpha \notin \mathbb{N}$	$\lfloor \alpha \rfloor$	$\frac{\alpha^{\lfloor \alpha \rfloor}}{\lfloor \alpha \rfloor!}e^{-\alpha}$

Pf. $\frac{p_{i+1}}{p_i} = \frac{\alpha}{i+1}$. The denominator is positive and ≥ 1. Therefore, for fixed α, $\frac{p_{i+1}}{p_i}$ is a non-increasing function of i. Hence for $\alpha < 1$, $\frac{p_{i+1}}{p_i} = \frac{\alpha}{i+1}$ always < 1, and p_i is always decreasing. For $\alpha \ge 1$, we search for $\frac{p_{i+1}}{p_i} = \frac{\alpha}{i+1} \le 1$ which yields $i \ge \alpha - 1$.

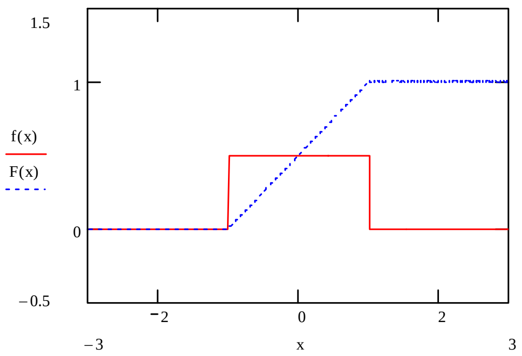
- peak value $p_{\lfloor \lambda \rfloor} \approx \frac{1}{\sqrt{2\pi\lambda}}$
- Rare events limit of the binomial (large n, small p)

  $\lim_{n\to\infty} p_i = \lim_{n\to\infty} \binom{n}{i} p^i (1-p)^{n-i} = e^{-\alpha} \frac{\alpha^i}{i!}$ where $\alpha = np$.
- Example
  - #photons emitted by a light source of intensity λ [photons/second] in time τ
  - #atoms of radioactive material undergoing decay in time τ
  - #clicks in a Geiger counter in τ seconds when the average number of click in 1 second is λ.
  - #dopant atoms deposited to make a small device such as an FET
  - #customers arriving in a queue or workstations requesting service from a file server in time τ
  - Counts of demands for telephone connections
  - number of occurrences of rare events in time τ
  - #soldiers kicked to death by horses
  - Counts of defects in a semiconductor chip.

PDF

Uniform: $\mathcal{U}(a,b)$

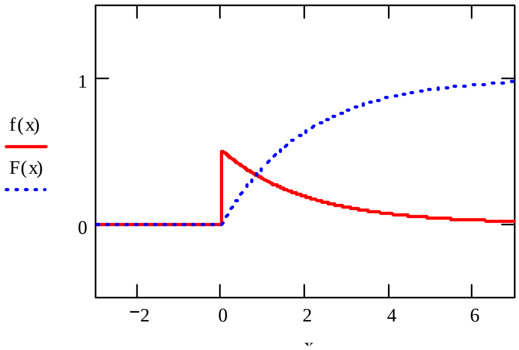
- $f(x) = \frac{1}{b-a} U(x-a) U(b-x) = \begin{cases} 0 & x < a, x > b \\ \frac{1}{b-a} & a \le x \le b \end{cases}$, $F(x) = \begin{cases} 0 & x < a, x > b \\ \frac{x-a}{b-a} & a \le x \le b \end{cases}$.



- $EX = \dfrac{a+b}{2}$
- $Var[X] = \dfrac{(b-a)^2}{12}$
- $\Phi_X(u) = e^{iu\frac{b+a}{2}} \dfrac{\sin\left(u\dfrac{b-a}{2}\right)}{u\dfrac{b-a}{2}}$
- continuous generalization of $\mathcal{R}(n)$
- use with caution to represent ignorance about a parameter taking value in [a,b]
- Ex.
  - phase of oscillators $\Rightarrow$ [-π,π] or [0,2π]
  - phase of received signals in incoherent communications $\rightarrow$ usual broadcast carrier phase $\phi \sim \mathcal{U}(-\pi,\pi)$
  - mobile cellular communication: multipath $\rightarrow$ path phases $\phi_c \sim \mathcal{U}(-\pi,\pi)$

Exponential: $\mathcal{E}(\lambda)$

- $f_X(x) = \lambda e^{-\lambda x} U(x)$; $\lambda > 0$
- $F_X(x) = \left(1 - e^{-\lambda x}\right) U(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$



- $\Pr[X > x] = e^{-\lambda x}u(x) + u(-x)$ i.e. $= \begin{cases} e^{-\lambda x}, & x \ge 0 \\ 1, & x < 0 \end{cases}$.
- Continuous version of $\mathcal{G}(\beta)$
- Exponential distribution $\mathcal{E}(\lambda)$ is $\Gamma(1,\lambda)$.
- Lack of memory property: $\Pr[X > k + c \mid X > k] = \Pr[X > c]$, for k, $c > 0$.

Pf. $\Pr[X > k + c \mid X > k] = \dfrac{\Pr[X > k + c \wedge X > k]}{\Pr[X > k]} = \dfrac{\Pr[X > k + c]}{\Pr[X > k]} = \dfrac{e^{-\alpha(k+c)}}{e^{-\alpha k}}$.

$$= e^{-\alpha c} = \Pr[X > c]$$

  - The future is independent of the past. The fact that it hasn't happened yet, tells us nothing about how much longer it will take before it does happen
- The exponential r.v. is the only continuous r.v. that satisfies the memoryless property.
- $\Phi_X(u) = \dfrac{\lambda}{\lambda - iu}$

Pf. $\Phi_X(u) = \mathrm{E}\left[e^{iuX}\right] = \int_{-\infty}^{\infty} e^{iux} f(x)dx = \int_{-\infty}^{\infty} e^{iux}\left(\lambda e^{-\lambda x}u(x)\right)dx$

$$= \lambda\int_0^{\infty} e^{-(\lambda - iu)x}dx = -\frac{\lambda}{\lambda - iu}e^{-(\lambda - iu)x}\Big|_0^{\infty} = \frac{\lambda}{\lambda - iu}$$

- $EX = \dfrac{1}{\lambda}$

Pf. $EX = \dfrac{\left(\dfrac{d}{du}\Phi_X(u)\right)\Big|_{u=0}}{i} = \dfrac{\left(-\dfrac{\lambda}{(\lambda - iu)^2}(-i)\right)\Big|_{u=0}}{i} = \dfrac{1}{\lambda}$

- $VAR(X) = \dfrac{1}{\lambda^2}$

Pf. $EX^2 = \dfrac{\left.\left(\dfrac{d^2}{du^2}\Phi_X(u)\right)\right|_{u=0}}{i^2} = -\left.\left((-2)\dfrac{\lambda i}{(\lambda - iu)^3}(-i)\right)\right|_{u=0} = \dfrac{2}{\lambda^2}$

$$\sigma_X^2 = EX^2 - (EX)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

- Ex
  - lifetimes (continuous time) of components of systems that fail without aging (memorylessness - eg wine glass)
  - waiting times between successive
    - photon arrivals
    - electron emissions from a cathode
    - radioactive decays
    - customer/packet arrivals
    - dopant atoms arrival in an implant process
  - duration of telephone or wireless call

- Independent $X_i \sim \mathcal{E}(\lambda_i)$. Y is the minimum of these X_i. (= Time till the first occurrence, if X_i denote time for each one to occur.)

  Then, Y ~ $\mathcal{E}(\beta)$; $\beta = \sum_{i=1}^{n} \lambda_i$.

Pf. $F_Y(y) = \Pr[Y \le y] = \Pr[X_1 \le y \vee X_2 \le y \vee \ldots \vee X_n \le y]$

$$= 1 - \Pr[X_1 > y, X_2 > y, \ldots, X_n > y] = 1 - \prod_{i=1}^{n} \Pr[X_i > y] = 1 - \prod_{i=1}^{n}\left(1 - F_X(y)\right)$$

$$= 1 - \prod_{i=1}^{n}\left(1 - \left(1 - e^{-\lambda_i y}\right)\right) = 1 - \prod_{i=1}^{n} e^{-\lambda_i y} \quad ; y > 0$$

$$= 1 - e^{-\left(\sum_{i=1}^{n}\lambda_i\right)y} \quad ; y > 0$$

- Half life: $P(T < h) = P(T \le h) = \dfrac{1}{2} \Rightarrow h = \dfrac{\ln 2}{\lambda}$

  Pf. $P(T \le h) = 1 - e^{-\lambda h} = \dfrac{1}{2} \Rightarrow e^{-\lambda h} = \dfrac{1}{2} \Rightarrow h = \dfrac{\ln 2}{\lambda}$

- Simulation of an exponential random variable from $U \sim \mathcal{U}(0,1)$

  To generate X whose $F_X(x)$ = $\left(1 - e^{-\lambda x}\right)u(x)$ from $U \sim \mathcal{U}(0,1)$

  Use, $X = -\dfrac{1}{\lambda}\ln U$

Pf. Let $G = F_X^{-1}(U)$. Note that for $F_X(x) = \left(1-e^{-\lambda x}\right)u(x)$, $F_X^{-1}(u):(0,1)\xrightarrow[\text{onto}]{1-1}(0,\infty)$.

$\Pr[G \le x] = \Pr\left[F_X^{-1}(U) \le x\right] = \Pr\left[U \le F_X(x)\right]$

Since $U \sim \mathcal{U}(0,1)$, $\Pr[U \le u] = u$.

Thus, $\Pr[G \le x] = \Pr\left[U \le F_X(x)\right] = F_X(x)$.

Hence, to generate X whose $F_X(x) = \left(1-e^{-\lambda x}\right)u(x)$ from $U \sim \mathcal{U}(0,1)$

Solving $1-e^{-\lambda X} = U$ for X yields $X = -\frac{1}{\lambda}\ln(1-U)$. Since $1-U$ is also uniformly distributed in [0,1], we can use the simpler expression $X = -\frac{1}{\lambda}\ln U$.

- N independent sources into a system. Waiting time between arrivals for each source ~ $\mathcal{E}(\lambda_i)$. Then
  - Waiting time between arrivals for system ~ $\mathcal{E}(\Sigma\lambda_i)$
  - Given that a customer arrives at the system, Pr[this customer is from source k] = $\frac{\lambda_k}{\sum_{i=1}^{N}\lambda_i}$

- Pf. By memoryless property, time between arrivals for system is the minimum of the time till next arrival for all sources.
- Pf. Consider at $t = 0$. Let T_i = time from 0 till customer from source i arrive. Then $T_i \sim \mathcal{E}(\lambda_i)$.

$$\Pr[\text{next customer is from } T_k] = \Pr\left[\bigcap_{\substack{i=1\\i\ne k}}^{N}[T_i > T_k]\right] = \int_0^\infty f_{T_k}(t)\prod_{i\ne k}P[T_i > t]dt$$

$$= \int_0^\infty f_{T_k}(t)\prod_{i\ne k}\left(1-F_{T_i}(t)\right)dt = \int_0^\infty \lambda_k e^{-\lambda_k t}\prod_{i\ne k}e^{-\lambda_i t}dt$$

$$= \int_0^\infty \lambda_k\prod_{i=1}^{N}e^{-\lambda_i t}dt = \lambda_k\int_0^\infty e^{-\left(\sum_{i=1}^{N}\lambda_i\right)t}dt = \frac{\lambda_k}{\sum_{i=1}^{N}\lambda_i}$$

- For S_i, T_j iid Exp(α), $P\left(\sum_{i=1}^{m}S_i > \sum_{j=1}^{n}T_j\right) = \sum_{i=0}^{m-1}\binom{n+m-1}{i}\left(\frac{1}{2}\right)^{n+m-1} = \sum_{i=0}^{m-1}\binom{n+i-1}{i}\left(\frac{1}{2}\right)^{n+i}$.

Pf. Method 1

Let A and B generate messages A_i and B_j respectively. At destination same for A and B, time between arrival of A_{i+1} and A_i is T_i. Time between B_{j+1} and B_j is S_j.

The set of times at which a message arrive is a Poisson process and

a point in this process is equally likely to be an arrival of messages from A or B, with independence between successive arrival.

Arrival at destination could be listed. Ex. A_1, A_2, B_1, B_2, A_3, A_4, B_3.

To have $\sum_{i=1}^{m} S_i > \sum_{j=1}^{n} T_j$, we want to have B_m arrive after A_n.

Equivalently, we need to create a list similar to the example above but have B_m after A_n.

A_n has to come after $A_1, A_2, \ldots, A_{n-1}$, and

B_m has to come after $B_1, B_2, \ldots, B_{m-1}$.

So, B_m has to come after $A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_{m-1}$.

This means that B_m is not among the first $n+m-1$ messages that arrive at destination.

Note that to have B_m arrive after A_n, the first $n+m-1$ messages don't have to have all $B_1, B_2, \ldots, B_{m-1}$.

We only need to have A_n and not B_m in the first n+m-1 messages.

In fact, we can have any amount of B's messages from 0 to m-1. If we have less than m-1 B's messages arrive at destination for the first $n+m-1$ messages, then we have A_i's, $i = n+1, n+2, \ldots$, in the first $n+m-1$ messages. Since B_m has to be after these, it is guaranteed to be after A_n.

So, we consider m-1 cases. Where each case denotes numbers of B's in the first n+m-1 messages arrived at the destination.

If we have i B's in the first $n+m-1$ messages,

then this i B's can choose their places in the $n+m-1$ position.

There are $\binom{n+m-1}{i}$ possibilities.

Then, A_i will fill up the rest of the positions.

However, each position can be A or B with equal probability.

The probability that it arrives according to one specific arrangement = $\left(\frac{1}{2}\right)^{n+m-1}$.

Method 2

Consider $i = 0$ to m-1 For each i, find all combination of A_1 to A_{n-1} with the first i B's. There are $\binom{n+i-1}{i}$ distinct sequences because we have to choose i positions for the B's from n-1+i positions. Now add A_n to the end of each sequence, yielding sequence of length $n+i$ which happens with probability $\left(\frac{1}{2}\right)^{n+i}$. The later part of the sequence is irrelevant. Because B_m has not yet been used in any of the sequence, each completed sequence will surely have B_m after A_n. Hence, for each i, the probability that B_m comes after A_n is $\binom{n+i-1}{i}\left(\frac{1}{2}\right)^{n+i}$.

Method 3

Let $S = \sum_{i=1}^{m} S_i$ and $T = \sum_{j=1}^{n} T_j$. Note that S and T are m-Erlang and n-Erlang r.v. respectively.

$$\Pr\left[S > T\right] = 1 - \Pr\left[S \le T\right] = \int_0^\infty \left(1 - F_S\left(t\right)\right) f_T\left(t\right) dt$$

$$= \int_0^\infty \sum_{k=0}^{m-1} \frac{\left(\lambda t\right)^k}{k!} e^{-\lambda t} \frac{\lambda^n t^{n-1} e^{-\lambda t}}{\left(n-1\right)!} dt = \frac{\lambda^n}{\left(n-1\right)!} \sum_{k=0}^{m-1} \frac{\lambda^k}{k!} \int_0^\infty t^{k+n-1} e^{-(2\lambda)t} dt$$

$$= \frac{\lambda^n}{\left(n-1\right)!} \sum_{k=0}^{m-1} \frac{\lambda^k}{k!} \frac{\left(k+n-1\right)!}{2^{k+n-1+1} \lambda^{k+n-1+1}} = \sum_{k=0}^{m-1} \binom{k+n-1}{k} \frac{1}{2^{k+n}}$$

Method 4 (limited) Exp(1)

- m = 1: $\Pr\left[S_1 > \sum_{j=1}^{n} T_j\right] = \sum_{i=0}^{1-1} \binom{n+1-1}{i} \left(\frac{1}{2}\right)^{n+1-1} = \left(\frac{1}{2}\right)^n$

$$\Pr\left[S_1 > \sum_{j=1}^{n} T_n\right] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(1 - F_S\left(\sum_{j=1}^{n} t_n\right)\right) \prod_{k=1}^{n} f_T\left(t_k\right) dt_1 \cdots dt_n$$

$$= \int_0^\infty \cdots \int_0^\infty e^{-\lambda \sum_{j=1}^{n} t_n} \prod_{k=1}^{n} \lambda e^{-\lambda t_k} dt_1 \cdots dt_n = \int_0^\infty \cdots \int_0^\infty e^{-\lambda \sum_{j=1}^{n} t_n} \lambda^n e^{-\lambda \sum_{k=1}^{n} t_k} dt_1 \cdots dt_n$$

$$= \int_0^\infty \cdots \int_0^\infty \lambda^n e^{-\sum_{j=1}^{n} 2\lambda t_n} dt_1 \cdots dt_n = \prod_{k=1}^{n} \left(\lambda \int_0^\infty e^{-2\lambda t_k} dt_k\right)$$

$$= \prod_{k=1}^{n} \left(-\frac{1}{2} e^{-2\lambda t_k} \Big|_0^\infty dt_k\right) = \prod_{k=1}^{n} \frac{1}{2} = \left(\frac{1}{2}\right)^n$$

- m = 2: $\Pr\left[S_1 + S_2 > \sum_{j=1}^{n} T_j\right] = \sum_{i=0}^{2-1} \binom{n+2-1}{i} \left(\frac{1}{2}\right)^{n+2-1} = \left(\binom{n+1}{0} + \binom{n+1}{1}\right)\left(\frac{1}{2}\right)^{n+1}$

$$= \left(n+2\right)\left(\frac{1}{2}\right)^{n+1}$$

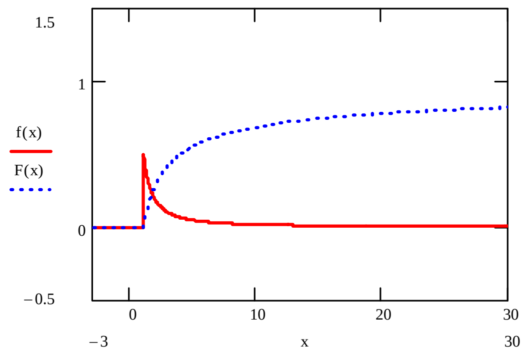
$$\Pr\left[S_1+S_2>\sum_{j=1}^{n}T_n\right]=\Pr\left[\left\{S_1+S_2>\sum_{j=1}^{n}T_n\right\}\wedge\left\{S_1>\sum_{j=1}^{n}T_n\right\}\right]$$
$$+\Pr\left[\left\{S_1+S_2>\sum_{j=1}^{n}T_n\right\}\wedge\left\{S_1\le\sum_{j=1}^{n}T_n\right\}\right]$$
$$=\Pr\left[S_1>\sum_{j=1}^{n}T_n\right]+\Pr\left[\left\{S_1+S_2>\sum_{j=1}^{n}T_n\right\}\wedge\left\{S_1\le\sum_{j=1}^{n}T_n\right\}\right]$$

We already know that $P\left(S_1>\sum_{j=1}^{n}T_n\right)=\left(\frac{1}{2}\right)^n$.

$$\Pr\left[\left\{S_1+S_2>\sum_{j=1}^{n}T_n\right\}\wedge\left\{S_1\le\sum_{j=1}^{n}T_n\right\}\right]$$
$$=\Pr\left[\left\{S_2>\left(\sum_{j=1}^{n}T_n\right)-S_1\right\}\wedge\left\{S_1<\sum_{j=1}^{n}T_n\right\}\right]$$
$$=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-\infty}^{\sum_{j=1}^{n}t_n}\left(1-F_{S_2}\left(\sum_{j=1}^{n}t_n\right)\right)f_{S_1}(s_1)\prod_{k=1}^{n}f_T(t_k)ds_1dt_1\cdots dt_n$$
$$=\int_{0}^{\infty}\cdots\int_{0}^{\infty}\int_{0}^{\sum_{j=1}^{n}t_n}e^{-\alpha\left(\left(\sum_{j=1}^{n}t_n\right)-s_1\right)}\alpha e^{-\alpha s_1}\alpha^n e^{-\alpha\sum_{j=1}^{n}t_n}ds_1dt_1\cdots dt_n$$
$$=\alpha^{n+1}\int_{0}^{\infty}\cdots\int_{0}^{\infty}e^{-2\alpha\left(\sum_{j=1}^{n}t_n\right)}\left(\sum_{j=1}^{n}t_n\right)dt_1\cdots dt_n$$
$$=\alpha^{n+1}\frac{n}{(2\alpha)^{n+1}}=\frac{n}{2^{n+1}}$$

Pareto: $\mathscr{Par}(\alpha)$:heavy-tailed model/density

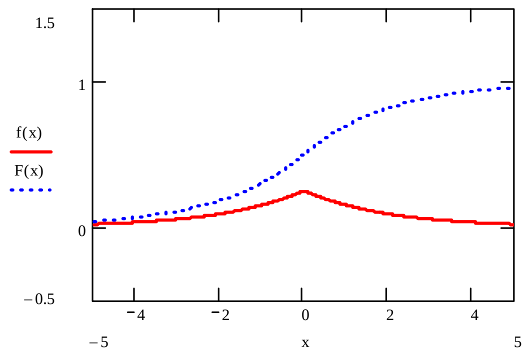
- $f(x)=\alpha x^{-\alpha-1}U(x-1);\ \alpha>0;\ F(x)=\left(1-\frac{1}{x^\alpha}\right)U(x-1)=\begin{cases}0 & x<1\\ 1-\frac{1}{x^\alpha} & x\ge 1\end{cases}$



- Ex
  - distribution of wealth
  - flood heights of the Nile river
  - designing dam height
  - (discrete) sizes of files requested by web users
  - waiting times between successive keystrokes at computer terminals
  - (discrete) sizes of files stored on Unix system file servers
  - running times for NP-hard problems as a function of certain parameters

Laplacian: $\mathcal{L}(\alpha)$

- $f(x) = \frac{\alpha}{2} e^{-\alpha|x|}$, $\alpha > 0$; $F(x) = \begin{cases} \frac{1}{2} e^{\alpha x} & x < 0 \\ 1 - \frac{1}{2} e^{-\alpha x} & x \geq 0 \end{cases}$



- $EX = 0$
- $Var[X] = \frac{2}{\alpha^2}$
- $\Phi_X(u) = \frac{\alpha^2}{\alpha^2 + u^2}$
- Ex
  - amplitudes of speech signals

- amplitudes of differences of intensities between adjacent pixels in an image

Normal/ Gaussian: $\mathcal{N}(m,\sigma^2)$

- $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
- $E\left[e^{-jvX}\right] = e^{jmv - \frac{1}{2}v^2\sigma^2}$.
- $\mathcal{N}\left(\vec{\mu}_{\vec{\theta}}, \vec{\Sigma}_{\vec{\theta}}\right) \; \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\Lambda)}} e^{-\frac{1}{2}(\underline{x}-\underline{m})^T \Lambda^{-1}(\underline{x}-\underline{m})}$; i.i.d. $\frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}} \exp\left\{-\frac{\|x_i-\mu\|^2}{2\sigma^2}\right\} =$
  $\frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{\mu^2 n}{2\sigma^2}\right\}$.
- $\mathcal{CN}\left(\vec{\mu}_{\vec{\theta}}, \vec{\Sigma}_{\vec{\theta}}\right) \cdot \frac{1}{\pi^n \det(\Lambda)} e^{-(\underline{x}-\underline{m})^H \Lambda^{-1}(\underline{x}-\underline{m})}$.
- $Q(0) = \frac{1}{2}$, $Q(-z) = 1 - Q(z)$ | $Q^{-1}(1-Q(z)) = -z$ | $P[X > x] = Q\left(\frac{x-m}{\sigma}\right)$
  | $P[X < x] = 1 - Q\left(\frac{x-m}{\sigma}\right) = Q\left(-\frac{x-m}{\sigma}\right)$

Rayleigh

- $F(x) = \left(1 - e^{-\alpha x^2}\right)u(x)$; $f(x) = 2\alpha x e^{-\alpha x^2} u(x)$
- $\Pr[X > t] = 1 - F(t) = \begin{cases} e^{-\alpha t^2} & t \ge 0 \\ 1 & t < 0 \end{cases}$
- Let X be a Rayleigh r.v., then $Y = X^2$ is an exponential r.v.

  Pf. $f_Y(y) = \begin{cases} 0, & y < 0 \\ \frac{1}{2\sqrt{y}} f_X\left(\sqrt{y}\right) + \frac{1}{2\sqrt{y}} f_X\left(-\sqrt{y}\right), & y \ge 0 \end{cases}$

  $= \begin{cases} 0, & y < 0 \\ \frac{1}{2\sqrt{y}} 2\alpha\sqrt{y} e^{-\alpha y} u\left(\sqrt{y}\right) - \cancelto{0}{\frac{1}{2\sqrt{y}} 2\alpha\sqrt{y} e^{-\alpha y} u\left(-\sqrt{y}\right)}, & y \ge 0 \end{cases} = \begin{cases} 0, & y < 0 \\ \alpha e^{-\alpha y}, & y \ge 0 \end{cases}$
- Ex
  - noise X at the output of AM envelope detector when no signal is present
- $X, Y \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma^2\right)$: If X and Y are independent, identically distributed, zero-mean normal random variables, then $R \equiv \sqrt{X^2 + Y^2}$ has a Rayleigh density.

- $X^2, Y^2 \overset{i.i.d.}{\sim} \Gamma\left(\frac{1}{2}, \frac{1}{2\sigma^2}\right)$. $X^2 + Y^2 \sim \Gamma\left(1, \alpha = \frac{1}{2\sigma^2}\right)$, exponential. $\sqrt{X^2 + Y^2}$ is a Rayleigh r.v. $\alpha = \frac{1}{2\sigma^2}$.
- Alternatively, transformation from Cartesian coordinates (x,y) to polar coordinates (r,θ) $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} r \\ \theta \end{pmatrix}$.

  $$f_{R,\Theta}(r,\theta) = r f_{X,Y}(r\cos\theta, r\sin\theta) = r\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{r\cos\theta}{\sigma}\right)^2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{r\sin\theta}{\sigma}\right)^2}$$
  $$= \left(\frac{1}{2\pi}\right)\left(2r\frac{1}{2\sigma^2} e^{-\frac{1}{2\sigma^2}r^2}\right)$$

  Hence, the radius R and the angle Θ are independent, with the radius R having a Rayleigh pdf while the angle Θ is uniformly distributed in the interval $(0, 2\pi)$.

Cauchy

- $f(z) = \frac{a}{\pi}\frac{1}{a^2 + z^2}$, a> 0.
- Mean and variance do not exist.
- $\Phi_X(u) = e^{-\alpha|u|}$.

Gamma distribution: $\Gamma(q, \lambda)$

- **Gamma distribution**: $\Gamma(q,\lambda)$. $f_X(x) = \frac{\lambda^q x^{q-1} e^{-\lambda x}}{\Gamma(q)} = \boxed{\frac{\lambda(\lambda x)^{q-1} e^{-\lambda x}}{\Gamma(q)}} = \frac{\left(\frac{x}{\alpha}\right)^{q-1} e^{-\frac{x}{\alpha}}}{\alpha\Gamma(q)}$; λ, q > 0.

  $x \geq 0$ for $q \geq 1$, > 0 for q < 1.
- $\Phi_X(u) = \frac{1}{\left(1 - i\frac{u}{\lambda}\right)^q}$

  Pf. $\Phi_X(u) = \int_0^\infty e^{iux} \frac{\lambda^q x^{q-1} e^{-\lambda x}}{\Gamma(q)} dx = \frac{\lambda^q}{\Gamma(q)} \int_0^\infty e^{-(\lambda - iu)x} x^{q-1} dx = \frac{\lambda^q}{\Gamma(q)} \frac{\Gamma(q)}{(\lambda - iu)^q}$.
- $EX = \frac{q}{\lambda}$

  Pf. $\frac{d}{du}\Phi_X(u) = -q\left(1 - i\frac{u}{\lambda}\right)^{-q-1}(-i)\frac{1}{\lambda}$. $EX = -i\frac{d}{du}\Phi_X(u)\bigg|_{u=0} = \frac{q}{\lambda}$
- $EX^2 = \frac{1}{\lambda^2}\left(q^2 + q\right)$

Pf. $\frac{d^2}{du^2}\Phi_X(u) = q(q+1)\left(1-i\frac{u}{\lambda}\right)^{-q-2}(-i)^2\frac{1}{\lambda^2}$

$$EX^2 = -\frac{d^2}{du^2}\Phi_X(u)\bigg|_{u=0} = q(q+1)\frac{1}{\lambda^2} = \frac{1}{\lambda^2}\left(q^2+q\right)$$

- $Var[X] = \frac{q}{\lambda^2}$
- Exponential distribution $\mathcal{E}(\lambda)$ is $\Gamma(1,\lambda)$.
- Chi-square random variable with *k* degrees of freedom: $\Gamma\left(\frac{k}{2},\frac{1}{2}\right)$.
- The *m*-Erlang r.v. is obtained when *q* = *m*, a positive integer.
  $f_X(x) = \frac{\lambda^m x^{m-1}e^{-\lambda x}}{\Gamma(m)} = \frac{\lambda^m x^{m-1}e^{-\lambda x}}{(m-1)!}$.
  - Let $X_i \sim Exp(\lambda)$, $S = \sum_{i=1}^{m} X_i$ is an *m*-Erlang r.v.

    Pf. $F_S(s) = \Pr[S \le s] = \Pr\left[\sum_{i=1}^{m} X_i \le s\right]$. Let N_s be the Poisson r.v. for the number of event

    in s seconds. Note that $\sum_{i=1}^{m} X_i \le s \equiv N_s \ge m$.

    Thus, $F_S(s) = \Pr[N_s \ge m] = 1 - \Pr[N_s < m] = 1 - \sum_{k=0}^{m-1}\frac{(\lambda s)^k}{k!}e^{-\lambda s}$.

$$f_S(s) = \frac{d}{ds}F_S(s) = -\sum_{k=0}^{m-1} k\frac{(\lambda s)^{k-1}}{k!}\lambda e^{-\lambda s} + \sum_{k=0}^{m-1}\lambda\frac{(\lambda s)^k}{k!}e^{-\lambda s}$$

$$= \left(\lambda e^{-\lambda s}\right)\left(\sum_{k=0}^{m-1}\frac{(\lambda s)^k}{k!} - \sum_{k=1}^{m-1}\frac{(\lambda s)^{k-1}}{(k-1)!}\right) = \left(\lambda e^{-\lambda s}\right)\frac{(\lambda s)^{m-1}}{(m-1)!}$$

- Let $X_i \sim \Gamma(q_i,\lambda)$, independent. Then
  - $Z_1 = \frac{X_1}{X_1+X_2}$ and $Z_2 = X_1 + X_2$ are independent. $Z_1 = \frac{X_1}{X_1+X_2} \sim \beta_{q_1,q_2}(z)$.
    $Z_2 = X_1 + X_2 \sim \Gamma(q_1+q_2,\lambda)$. Note that $Z_1 \in (0,1)$.
  - $Z_2 = X_1 + X_2$ and $Z_3 = \frac{X_1}{X_2}$ are independent, Z_3 is a beta prime distribution with parameters (q_1,q_2).

Pf. $f_{X_i}(x_i) = \frac{\lambda^{q_i} x_i^{q_i-1}e^{-\lambda x_i}}{\Gamma(q_i)}$. By independence,

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2) = \frac{\lambda^{q_1} x_1^{q_1-1} e^{-\lambda x_1}}{\Gamma(q_1)} \frac{\lambda^{q_2} x_2^{q_2-1} e^{-\lambda x_2}}{\Gamma(q_2)} = \frac{\lambda^{q_1+q_2} x_1^{q_1-1} x_2^{q_2-1} e^{-\lambda x_1} e^{-\lambda x_2}}{\Gamma(q_1)\Gamma(q_2)}.$$

Let $U = X_1 + X_2$ and $V = \dfrac{X_1}{X_1 + X_2}$. Then, $X_1 = UV$ and $X_2 = U - UV = U(1-V)$.

$$|\det J| = \left|\det\left(\begin{bmatrix} \frac{\partial}{\partial u} uv & \frac{\partial}{\partial v} uv \\ \frac{\partial}{\partial u} u(1-v) & \frac{\partial}{\partial v} u(1-v) \end{bmatrix}\right)\right| = \left|\det\left(\begin{bmatrix} v & u \\ 1-v & -u \end{bmatrix}\right)\right|$$

$$= |-vu - u + uv| = |-u| = u$$

Hence, $f_{U,V}(u,v) = f_{X,Y}(uv, u(1-v))|\det J|$

$$= \frac{\lambda^{q_1+q_2} (uv)^{q_1-1} (u(1-v))^{q_2-1} e^{-\lambda uv} e^{-\lambda u(1-v)}}{\Gamma(q_1)\Gamma(q_2)} u$$

$$= \frac{\lambda^{q_1+q_2} u^{q_1+q_2-1} e^{-\lambda u}}{\Gamma(q_1+q_2)} \frac{\Gamma(q_1+q_2)}{\Gamma(q_1)\Gamma(q_2)} v^{q_1-1} (1-v)^{q_2-1}$$

Pf. $U = X_1 + X_2$ and $W = \dfrac{X_1}{X_2}$. Then $X_1 = U\dfrac{W}{W+1}$, $X_2 = \dfrac{U}{W+1}$.

$$|\det J| = \left|\det\left(\begin{bmatrix} \frac{\partial}{\partial u} u\frac{w}{w+1} & \frac{\partial}{\partial w} u\frac{w}{w+1} \\ \frac{\partial}{\partial u} \frac{u}{w+1} & \frac{\partial}{\partial w} \frac{u}{w+1} \end{bmatrix}\right)\right| = \left|\det\left(\begin{bmatrix} \frac{w}{w+1} & u\frac{(w+1)-w}{(w+1)^2} \\ \frac{1}{w+1} & \frac{-u}{(w+1)^2} \end{bmatrix}\right)\right|$$

$$= \left|\frac{w}{w+1}\frac{u}{(w+1)^2} + u\frac{1}{(w+1)^2}\frac{1}{w+1}\right| = \left|\frac{u}{(w+1)^2}\right| = \frac{u}{(w+1)^2}$$

$$f_{U,W}(u,w) = f_{X,Y}\left(u\frac{w}{w+1}, \frac{u}{w+1}\right)|\det J|$$

$$= \frac{\lambda^{q_1+q_2}\left(u\frac{w}{w+1}\right)^{q_1-1}\left(u\frac{1}{w+1}\right)^{q_2-1} e^{-\lambda u \frac{w}{w+1}} e^{-\lambda u \frac{1}{w+1}}}{\Gamma(q_1)\Gamma(q_2)} \frac{u}{(w+1)^2}$$

$$= \frac{\lambda^{q_1+q_2} u^{q_1+q_2-1} e^{-\lambda u}}{\Gamma(q_1+q_2)} \frac{\Gamma(q_1+q_2)}{\Gamma(q_1)\Gamma(q_2)} w^{q_1-1} (w+1)^{-(q_1+q_2)}$$

- By induction, $\sum_i X_i \sim \Gamma\left(\sum_i q_i, \lambda\right)$.

- Ex.
  - The time required to service customers in queueing systems
  - Lifetime of devices and systems in reliability studies

- Defect clustering behavior in VLSI chips.

Beta distribution

- Beta distribution: $p(z) = \beta_{q_1,q_2}(z) = \frac{\Gamma(q_1+q_2)}{\Gamma(q_1)\Gamma(q_2)} z^{q_1-1}(1-z)^{q_2-1};\ z \in (0,1)$

Beta prime distribution

- $f_X(x) = \frac{\Gamma(q_1+q_2)}{\Gamma(q_1)\Gamma(q_2)} x^{q_1-1}(x+1)^{-(q_1+q_2)}$. $X > 0$.
- Let $X_i \sim \Gamma(q_i,\lambda)$, independent. Then $\frac{X_1}{X_2}$ is a beta prime distribution with parameters (q_1,q_2).
  - Let $X_i \sim \mathcal{E}(\lambda)$, independent. Then $X = \frac{X_1}{X_2}$ is a beta prime distribution with parameters $(1,1)$. $f_X(x) = \frac{1}{(x+1)^2}$, $x > 0$.

Chi-sqaure

- $X \sim \mathcal{N}(0,\sigma^2)$. Then $Y = X^2 \sim \Gamma\left(\frac{1}{2},\frac{1}{2\sigma^2}\right)$. Then $p(y) = \frac{1}{\sqrt{2\pi y}\sigma} e^{-\frac{y}{2\sigma^2}}$, $y \geq 0$.

  $\Phi(u) = \frac{1}{(1-j2u\sigma^2)^{\frac{1}{2}}}$.

  Pf. $f_Y(y) = \begin{cases} 0, & y<0 \\ \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}), & y \geq 0 \end{cases}$

  $= \begin{cases} 0, & y<0 \\ \frac{1}{2\sqrt{y}}\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{y}{\sigma^2}} + \frac{1}{2\sqrt{y}}\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{y}{\sigma^2}}, & y \geq 0 \end{cases} = \begin{cases} 0, & y<0 \\ \frac{1}{\sqrt{2\pi y}\sigma} e^{-\frac{1}{2}\frac{y}{\sigma^2}}, & y \geq 0 \end{cases}$

  Note that $\frac{1}{\sqrt{2\pi y}\sigma} e^{-\frac{1}{2\sigma^2}y} = \frac{1}{\Gamma\left(\frac{1}{2}\right)} y^{\frac{1}{2}-1}\left(\frac{1}{2\sigma^2}\right)^{\frac{1}{2}} e^{-\frac{1}{2\sigma^2}y}$.

- Let $X \sim \mathcal{N}(\mu,\sigma^2)$: $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$. $Y = (X-\mu)^2 \sim \Gamma\left(\frac{1}{2},\frac{1}{2\sigma^2}\right)$

- **Chi-square**: $X_i \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma^2)$. $Y = \sum_{i=1}^{n} X_i^2$. Then

  $p(y) = \frac{1}{(2\sigma^2)^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2\sigma^2}} = \Gamma\left(\frac{n}{2},\frac{1}{2\sigma^2}\right), y \geq 0.$ $\Phi(u) = (1 - j2u\sigma^2)^{-\frac{n}{2}}$. $\mathbb{E}[Y] = n\sigma^2$,

  $Var[Y] = 2n\sigma^4$.

  Pf. We know that $X_i^2 \sim \text{iid } \Gamma\left(\frac{1}{2},\frac{1}{2\sigma^2}\right)$. Hence, $\sum_{i=1}^{n} X_i^2 \sim \Gamma\left(\frac{n}{2},\frac{1}{2\sigma^2}\right)$.

- Chi-square random variable with *k* degrees of freedom: $\Gamma\left(\frac{k}{2},\frac{1}{2}\right)$

  - The sum of *k* mutually independent, squared zero-mean unit-variance Gaussian random variables.

  - $f_X(x) = \frac{\left(\frac{1}{2}x\right)^{\frac{k}{2}-1} e^{-\frac{1}{2}x}}{2\Gamma\left(\frac{k}{2}\right)}$

Student's *t*-distribution

- $f_V(v) = \frac{\left(1+\frac{v^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}$

- Let *X* be a 0-mean, unit-variance Gaussian r.v., and let *Y* be a chi-square r.v. with *n* degrees of freedom. Assume that *X* and *Y* are independent. Let $V = \frac{X}{\sqrt{\frac{Y}{n}}}$. $W = Y$. Then $X = V\sqrt{\frac{W}{n}}$.

$$\det\begin{bmatrix} \frac{\partial}{\partial v} v\sqrt{\frac{w}{n}} & \frac{\partial}{\partial w} v\sqrt{\frac{w}{n}} \\ \frac{\partial}{\partial v} w & \frac{\partial}{\partial w} w \end{bmatrix} = \det\begin{bmatrix} \sqrt{\frac{w}{n}} & \frac{v}{2\sqrt{wn}} \\ 0 & 1 \end{bmatrix} = \sqrt{\frac{w}{n}}. \quad f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \frac{\left(\frac{1}{2}y\right)^{\frac{n}{2}-1} e^{-\frac{1}{2}y}}{2\Gamma\left(\frac{y}{2}\right)}.$$

$$f_V(v)=\int_0^\infty f_{V,W}(v,w)dw=\frac{1}{2^{\frac{n}{2}-1+1+\frac{1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_0^\infty w^{\frac{n-1}{2}}e^{-\frac{1}{2}\left(1+\frac{v^2}{n}\right)w}dw$$

$$=\frac{1}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_0^\infty w^{\frac{n+1}{2}-1}e^{-\frac{1}{2}\left(1+\frac{v^2}{n}\right)w}dw$$

$$=\frac{1}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\frac{1}{2^{\frac{n+1}{2}}}\left(1+\frac{v^2}{n}\right)^{\frac{n+1}{2}}}=\frac{1}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left(1+\frac{v^2}{n}\right)^{\frac{n+1}{2}}}$$

Binomial, Geometric, Exponential, Poisson

- Suppose *N*, the number of event occurrences in a *T*-second time interval is being counted. Divide the time interval into a very large number, *n*, of subintervals of length *T*/*n*. Each subinterval can be viewed as a Bernoulli trial with probability of success $p = \alpha/n$, if the following conditions hold:
  1) At most one event can occur in a subinterval, that is, the probability of more than one event occurrence is negligible.
  2) The outcomes in different subintervals are independent
  3) The probability of an event occurrence in a subinterval is $p=\frac{\alpha}{n}$, where α is the average number of events observed in a *T*-second interval.
- N ~ binomial $\mathcal{B}\left(n,\frac{\alpha}{n}\right)$.
- α is fixed.
- Rare events limit of the binomial (large *n*, small *p*)

  $\lim_{n\to\infty} p_i = \lim_{n\to\infty}\binom{n}{i}p^i(1-p)^{n-i} = e^{-\alpha}\frac{\alpha^i}{i!}$ where $\alpha = np$.

  Pf. Keep $\alpha = np$ fixed. Then $\lim_{n\to\infty}\binom{n}{0}p^0(1-p)^{n-0}=\lim_{n\to\infty}(1-p)^n=\lim_{n\to\infty}\left(1-\frac{\alpha}{n}\right)^n=e^{-\alpha}$.

  $$\frac{p_{i+1}}{p_i}=\frac{\binom{n}{i+1}p^{i+1}(1-p)^{n-i-1}}{\binom{n}{i}p^i(1-p)^{n-i}}=\frac{(n-i)p}{(i+1)(1-p)}=\frac{\left(1-\frac{i}{n}\right)np}{(i+1)\left(1-\frac{\alpha}{n}\right)}.$$

  Hence, $\lim_{n\to\infty}\frac{p_{i+1}}{p_i}=\frac{np}{(i+1)}=\frac{\alpha}{i+1}$.

Thus, the limiting probabilities satisfy $p_0 = e^{-\alpha}$ and $p_{i+1} = \frac{\alpha}{i+1} p_i$. By induction, this is the Poisson pmf.

- The number of subintervals M until (but not including) the occurrence of an event is a geometric r.v. (Still a geometric r.v. but different one if include the event's subinterval.)
- The time until the occurrence of the first event is $X = M\frac{T}{n}$.

$$\Pr\left[X > t\right] = \Pr\left[M \ge n\frac{t}{T}\right] = \left(1-p\right)^{n\frac{t}{T}} = \left(\left(1-\frac{\alpha}{n}\right)^n\right)^{\frac{t}{T}}$$

$$\lim_{n\to\infty}\Pr\left[X > t\right] = e^{-\alpha\frac{t}{T}}.$$

Thus, the exponential r.v. is obtained as a limiting form of the geometric random variable.

- This result implies that for a Poisson random variable, the time between events is an exponentially distributed random variable with parameter $\lambda = \frac{\alpha}{T}$ [events/sec]

Etc.

- Exponential and normal distribution: $X_1, X_2 \overset{iid}{\sim} \mathcal{N}\left(0,\sigma^2\right)$. Then $X_1^2 + X_2^2 \sim Exp\left(\frac{1}{2\sigma^2}\right)$.

Pf. $X_1^2 + X_2^2 \sim \Gamma\left(\frac{2}{2}, \frac{1}{2\sigma^2}\right) = Exp\left(\frac{1}{2\sigma^2}\right)$.

Example

- The total number of defects X on a chip is a Poisson r.v. with mean α. Suppose that each defect has a probability p of falling in a specific region R and that the location of each defect is independent of the locations of all other defects. Find the pmf of the number of defects Y that fall in the region R.

Solution:

We have $P\left[X = k\right] = e^{-\alpha}\frac{\alpha^k}{k!}$, and $P\left[Y = j \middle| X = k\right] = \binom{k}{j} p^j \left(1-p\right)^{k-j}$ for $0 \le j \le k$. Note that for $k < j$, $P\left[Y = j \middle| X = k\right] = 0$. Hence, from $P\left[Y = j\right] = \sum_{k=0}^{\infty} P\left[Y = j \middle| X = k\right] P\left[X = k\right]$, we have

$$P[Y=j]=\sum_{k=j}^{\infty}\binom{k}{j}p^j(1-p)^{k-j}e^{-\alpha}\frac{\alpha^k}{k!}=\sum_{k=j}^{\infty}\frac{k!}{j!(k-j)!}p^j(1-p)^{k-j}e^{-\alpha}\frac{\alpha^{k-j}}{k!}\alpha^j$$

$$=e^{-\alpha}\frac{(\alpha p)^j}{j!}\sum_{k=j}^{\infty}\frac{(\alpha(1-p))^{k-j}}{(k-j)!}\overset{\ell=k-j}{=}e^{-\alpha}\frac{(\alpha p)^j}{j!}\sum_{\ell=0}^{\infty}\frac{(\alpha(1-p))^{\ell}}{\ell!}$$

$$=e^{-\alpha}\frac{(\alpha p)^j}{j!}e^{\alpha(1-p)}=e^{-\alpha p}\frac{(\alpha p)^j}{j!}$$

Thus, Y is a Poisson r.v. with mean αp.

- The number of customers that arrive at a service station during a time interval t is a Poisson r.v. with parameter βt. The time T required to service each customer is an exponential r.v. with parameter α. Assume that the customer arrivals are independent of the customer service time. Find the pmf for the number of customers N that arrive during the service time T of a specific customer.

  Solution

  We have $P[T=t]=\alpha e^{-\alpha t}u(t)$, and $P\left[N=k\middle|T=t\right]=e^{-\beta t}\frac{(\beta t)^k}{k!}$. Hence,

$$P[N=k]=\int_0^{\infty}e^{-\beta t}\frac{(\beta t)^k}{k!}\alpha e^{-\alpha t}dt=\frac{\alpha\beta^k}{k!}\int_0^{\infty}e^{-(\beta+\alpha)t}t^k dt$$

$$=\frac{\alpha\beta^k}{k!}\frac{k!}{(\beta+\alpha)^{k+1}}=\frac{\alpha\beta^k}{(\beta+\alpha)^{k+1}}=\frac{\alpha}{\beta+\alpha}\left(\frac{\beta}{\beta+\alpha}\right)^k$$

- Let N = the random number of packets that arrive at A during the time interval [0,1]. Assume N is a Poisson random variable with mean λ. Each packet arriving at A is immediately sent to B with probability p or to C with probability q = 1-p, independently from packet to packet.

  Let L = number of packets sent to B during [0,1]

  R = number of packets sent to C during [0,1]

  Then,

  - L is a Poisson random variable with mean λp: $\Pr[L=\ell]=\frac{(\lambda p)^{\ell}e^{-\lambda p}}{\ell!}$

    Proof.

    First, note that $\Pr\left[L=\ell\middle|N=n\right]=\begin{cases}0 & \text{if } n<\ell\\ \binom{n}{\ell}p^{\ell}q^{n-\ell} & \text{if } n\geq\ell\end{cases}$

$$\Pr[L=\ell]=\sum_{n=0}^{\infty}\Pr[L=\ell|N=n]\Pr[N=n]$$

$$=\sum_{n=\ell}^{\infty}\left(\binom{n}{\ell}p^{\ell}q^{n-\ell}\right)\left(\frac{\lambda^n e^{-\lambda}}{n!}\right)=\sum_{n=\ell}^{\infty}\left(\frac{\cancel{n!}}{\ell!(n-\ell)!}p^{\ell}q^{n-\ell}\frac{\lambda^n e^{-\lambda}}{\cancel{n!}}\right)$$

$$=\sum_{m=0}^{\infty}\left(\frac{\lambda^{m+\ell}e^{-\lambda}}{\ell!(m)!}p^{\ell}q^{m}\right)=\frac{(\lambda p)^{\ell}e^{-\lambda}}{\ell!}\sum_{m=0}^{\infty}\frac{(\lambda q)^m}{m!}$$

$$=\frac{(\lambda p)^{\ell}e^{-\lambda}}{\ell!}e^{\lambda q}=\frac{(\lambda p)^{\ell}e^{-\lambda(1-q)}}{\ell!}=\frac{(\lambda p)^{\ell}e^{-\lambda p}}{\ell!}$$

- $\Pr[R=r]=\dfrac{(\lambda q)^r e^{-\lambda q}}{r!}$
- L and R are statistically independent

  Proof.

$$\Pr[L=\ell,R=r]=\Pr[L=\ell,N=\ell+r]=\Pr[L=\ell|N=\ell+r]\Pr[N=\ell+r]$$

$$=\left(\binom{\ell+r}{\ell}p^{\ell}q^{(\ell+r)-\ell}\right)\left(\frac{\lambda^{\ell+r}e^{-\lambda}}{(\ell+r)!}\right)=\frac{\cancel{(\ell+r)!}}{\ell!r!}p^{\ell}q^{r}\lambda^{\ell}\lambda^{r}\frac{e^{-\lambda}}{\cancel{(\ell+r)!}}$$

$$=\frac{(p\lambda)^{\ell}}{\ell!}\frac{(q\lambda)^{r}}{r!}e^{-\lambda}=\Pr[L=\ell]\cancel{e^{\lambda p}}\Pr[R=r]\cancel{e^{\lambda q}e^{-\lambda}}$$

Table

Model	p_i	
Random $\mathscr{R}_n$:	$\frac{1}{n}$; $\Omega=\{1,2,\ldots,n\}$	
Binomial $\mathscr{B}(n,p)$	$\binom{n}{k}p^k(1-p)^{n-k}$; Ω = N_{n+1} ; 0 ≤ p ≤1 • Bernoulli is $\mathscr{B}(1,p)$	
Geometric $\mathscr{G}(\beta)$	$(1-\beta)\beta^k$; Ω = N; 0 ≤ β < 1 k = #failures before (excluding) first success	
Poisson $\mathscr{P}(\lambda)$	$e^{-\lambda}\frac{\lambda^i}{i!}$; Ω = N, 0 ≤ λ	
Model	$f_X(x)$	$F_X(x)$
Uniform $\mathscr{U}(a,b)$	$\begin{cases}0 & x<a,x>b\\ \frac{1}{b-a} & a\le x\le b\end{cases}$	$\begin{cases}0 & x<a,x>b\\ \frac{x-a}{b-a} & a\le x\le b\end{cases}$
Exponential $\mathscr{E}(\alpha)$	$\alpha e^{-\alpha x}U(x)$; α > 0	$(1-e^{-\alpha x})U(x)$

Pareto $\mathcal{P}ar(\alpha)$	$\alpha x^{-\alpha-1}U(x-1)$; $\alpha > 0$	$\left(1-\frac{1}{x^{\alpha}}\right)U(x-1)$
Laplacian $\mathcal{L}(\alpha)$	$\frac{\alpha}{2}e^{-\alpha\|x\|}$; $\alpha > 0$	$\begin{cases} \frac{1}{2}e^{\alpha x} & x<0 \\ 1-\frac{1}{2}e^{-\alpha x} & x \geq 0 \end{cases}$
Normal $\mathcal{N}(m,\sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$; $\sigma > 0$	$erf\left(\frac{x-m}{\sigma}\right)$
Gamma distribution $\Gamma(q,\lambda)$	$\frac{\lambda^q x^{q-1}e^{-\lambda x}}{\Gamma(q)}$	

- Chi-square r.v. with k degrees of freedom is $\Gamma\left(\frac{k}{2},\frac{1}{2}\right)$.

Model	$E[X]$	$Var[X]$	$\Phi_X(x)$
Random $\mathcal{R}_n$	$\frac{n-1}{2}$	$\frac{n^2-1}{12}$	$\frac{1-e^{iun}}{n(1-e^{iu})}$
Binomial $\mathcal{B}(n,p)$	np	$np(1\text{-}p)$	$(1-p+pe^{ju})^n$
Geometric $\mathcal{G}(\beta)$	$\frac{\beta}{1-\beta}$	$\frac{\beta}{(1-\beta)^2}$	$\frac{1-\beta}{1-\beta e^{iu}}$
Poisson $\mathcal{P}(\lambda)$	λ	λ	$e^{\lambda(e^{iu}-1)}$
Uniform $\mathcal{U}(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$e^{iu\frac{b+a}{2}}\frac{\sin\left(u\frac{b-a}{2}\right)}{u\frac{b-a}{2}}$
Exponential $\mathcal{E}(\alpha)$	$\frac{1}{\alpha}$	$\frac{1}{\alpha^2}$	$\frac{\alpha}{\alpha-iu}$
Pareto $\mathcal{P}ar(\alpha)$	$\frac{\alpha}{\alpha-1}$, $\alpha>1$ ∞, $0<\alpha<1$	Undefined, $0<\alpha<1$ ∞, $1<\alpha<2$ $\frac{\alpha}{(\alpha-2)(\alpha-1)^2}$, $\alpha>2$	
Laplacian $\mathcal{L}(\alpha)$	0	$\frac{2}{\alpha^2}$	$\frac{\alpha^2}{\alpha^2+u^2}$
Normal/Gaussian $\mathcal{N}(m,\sigma^2)$	m	σ^2	$e^{ium-\frac{1}{2}\sigma^2u^2}$ $e^{i\underline{m}^T\underline{u}-\frac{1}{2}\underline{u}^TC\underline{u}}$
Gamma distribution $\Gamma(q,\lambda)$	$\frac{q}{\lambda}$	$\frac{q}{\lambda^2}$	$\frac{1}{\left(1-i\frac{u}{\lambda}\right)^q}$