Exponential random variables

- Exponential random variables arise in studies of waiting times, service times, etc
- **X** is called an exponential random variable with parameter λ if its pdf is given by
  - f(u) = λ•exp(–λu) for u ≥ 0
  - f(u) = 0 for u < 0
  - Scale parameter λ > 0
- E[**X**] = 1/λ   var(**X**) = $1/\lambda^2$

CDF of exponential RV

- **X** = exponential RV with parameter λ
- F(t) = P{**X** ≤ t} = area under pdf to left of t
  = 1 – exp(–λt)
- P{**X** > t} = exp(–λt) = complementary CDF
- E[**X**] = ∫P{**X** > t} dt = 1/λ
- P{**X** > t+τ | **X** > t} = P{**X** > t+τ}/P{**X** > t}
  = exp(–λτ)
- Memoryless property of exponential RVs

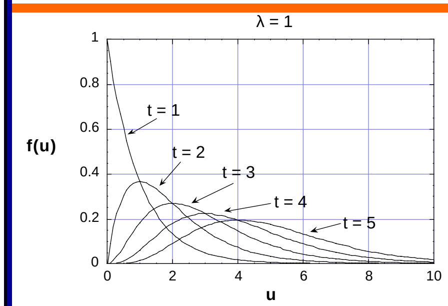
Gamma random variables

- **X** is called a gamma random variable with parameters (t, λ) if its pdf is given by
  - f(u) = λ•exp(–λu)•$(\lambda u)^{t-1}$/Γ(t) for u > 0
  - f(u) = 0 for u ≤ 0
  - t > 0 is the order parameter
  - λ > 0 is the scale parameter
  - Γ(t) is a number whose value is the gamma function evaluated at t

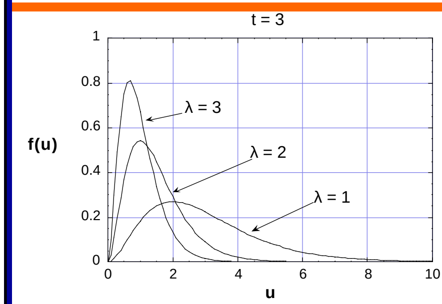
What's this Γ(t) stuff, anyway?

- The value of Γ(t) is given by
- $\Gamma(t) = \int_0^\infty x^{t-1} \cdot \exp(-x)\,dx, \quad t > 0$
- Γ(t) = (t–1)•Γ(t–1) = (t–1)•(t–2)•Γ(t–2) = …
- If t is an integer, Γ(t) = (t–1)!
- If t ≠ integer, Γ(t) = (t–1)•(t–2)•…•Γ(t–$\lfloor t \rfloor$) where t–$\lfloor t \rfloor$ is the fractional part of t
- For 0 < t < 1, numerical integration must be used to evaluate Γ(t)

Gamma pdfs: different values of t



Gamma pdfs: different values of λ



Mean & variance of gamma RVs

- If **X** is a gamma RV with parameters (t, λ), then E[**X**] = t/λ and var(**X**) = t/λ^2
- A gamma RV with order parameter t = 1 is an exponential RV with parameter λ
- A gamma RV with order parameter t = n is called an n-Erlang random variable
- A gamma RV with t = n/2, λ = 1/2 is a chi-square RV with n degrees of freedom

An analogy

- The exponential RV with parameter λ is analogous to the geometric RV with parameter p — both denote the waiting time till something occurs
- E[**X**] = 1/λ ↔ E[**X**] = 1/p
- The gamma RV with parameters (n, λ) is analogous to the negative binomial RV with parameters (n, p) — both denote the waiting time for the n-th occurrence
- E[**X**] = n/λ ↔ E[**X**] = n/p

More on the analogy

- With geometric and negative binomial RVs, we are, in effect, measuring the waiting time till something occurs in discrete steps
- With exponential and gamma RVs, we are modeling time as a continuous variable
- With negative binomial and gamma RV, the waiting time for the n-th occurrence of something is the sum of n (independent) waiting times for one thing to occur

What thing occurs?

- We are going to study occurrences of various interesting phenomena
  - telephone "off-hook" signals
  - jobs arriving at a processor
  - packets arriving at a router
  - α-particle emissions
  - gate failures in a TMR system
  - cars passing a checkpoint

What's common to all these?

- The occurrences of the phenomenon, e.g. a telephone going off-hook, a packet arriving at a router are not very frequent
- Actually, this infrequency depends on the time scale being used
- For a telephone, the chances that it will go off hook in the next millisecond are small
- P{off-hook at some time during next day} is close to 1

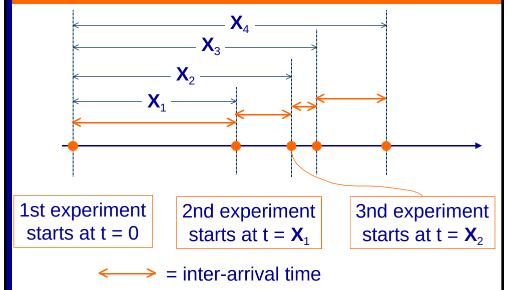
Occurrences at random

- The time interval between successive occurrences (called the inter-arrival time) varies at random
- The inter-arrival time is modeled as a random variable
- What is the probabilistic behavior of this random variable?
- Answer: geometric for discrete time
  exponential for continuous time

Independence of occurrences

- The observed value of a random variable depends on the outcome of an experiment
- An experiment commences at t = 0 and ends with the first arrival (i.e. occurrence) at some random time (say t = $\mathbf{X}_1$) later
- At this time, a new experiment is started, and it ends with the second arrival at some time t = $\mathbf{X}_2$, and so on…
- These experiments are independent

An illustrative diagram



Long term average = expectation

- The inter-arrival time is a random variable
- The successive inter-arrival times are the observed values of this random variable on independent trials
- Long-term observed average = (total time till N-th arrival occurs)/N is roughly equal to the expectation of the inter-arrival time

The arrival rate

- The inter-arrival time is a random variable with expected value (average) = 1/μ
- For large N, the total of N inter-arrival times is roughly N•(1/μ)
- The number of occurrences of the phenomenon of interest (arrivals) in a long time interval of duration T is roughly μT
- μ is called the arrival rate
- On average, μT arrivals in T seconds

Random arrivals in time

- We are considering some phenomenon that occurs at random times
- Sometimes, these occurrences are called events
- Here, event does not have its usual meaning in probability theory, ***viz.*** a subset of the sample space
- To avoid confusion, we shall call these occurrences as arrivals or points

Poisson process with arrival rate μ

- The sequence of random arrivals or points (with independent inter-arrival times having average value 1/μ) is called a Poisson process with arrival rate μ
- Why Poisson?
- Let $\mathbf{N}(t_1, t_2]$ denote the number of arrivals in the time interval $(t_1, t_2]$. Note endpoints
- $\mathbf{N}(t_1, t_2]$ is a Poisson random variable with parameter $\mu(t_2 - t_1)$ = μ•duration of interval

Basic assumptions

- At any time instant t, at most one arrival can occur
- Assumption: For small values of ΔT
  - $P\{\mathbf{N}(t, t+\Delta t] = 1\} \approx \mu \bullet \Delta T$
  - $P\{\mathbf{N}(t, t+\Delta t] = 0\} \approx 1 - \mu \bullet \Delta T$
- Approximation improves as $\Delta T \to 0$
- Approximation is nonsense if $\Delta T > 1/\mu$
- Small values of ΔT means $\Delta T << 1/\mu$

Is this consistent with Poissonity?

- We have not yet proved that $\mathbf{N}(t_1, t_2]$ is a Poisson RV with parameter $\mu \bullet (t_2 - t_1)$
- But, if $\mathbf{N}(t, t+\Delta t]$ were a Poisson random variable with parameter $\mu \bullet \Delta T << 1$, then

  $P\{\mathbf{N}(t, t+\Delta t] = 0\} = \exp(-\mu \bullet \Delta T)$

  $= 1 - \mu \bullet \Delta T + (\mu \bullet \Delta T)^2/2 - \ldots$

  $\approx 1 - \mu \bullet \Delta T$ for small $\mu \bullet \Delta T$

  $P\{\mathbf{N}(t, t+\Delta t] = 1\} = (\mu \bullet \Delta T) \approx \mu \bullet \Delta T$

The independence assumption

- Let $t_1 < t_2 \le t_3 < t_4 \le t_5 < t_6 \ldots$
- Independence assumption:
  - The random variables $\mathbf{N}(t_1, t_2]$, $\mathbf{N}(t_3, t_4]$, $\mathbf{N}(t_5, t_6]$, … are independent
  - The numbers of arrivals in disjoint intervals (i.e. non-overlapping intervals) of time are independent
  - Memoryless property of the process

Distribution of the first arrival time

- The first arrival occurs at random time $\mathbf{X}_1$
- We set up and solve a differential equation for $P_0(t) = P\{$no arrivals in $(0, t]\}$

  $= P\{\mathbf{N}(0, t] = 0\} = P\{\mathbf{X}_1 > t\}$
- $P_0(t+\Delta T) = P\{$no arrivals in $(0, t+\Delta T]\}$

  $= P\{\mathbf{N}(0, t+\Delta T] = 0\}$

  $= P[\{\mathbf{N}(0, t]=0\} \cap \{\mathbf{N}(t, t+\Delta T]=0\}]$

  $= P\{\mathbf{N}(0, t] = 0\} \bullet P\{\mathbf{N}(t, t+\Delta T] = 0\}$

The differential equation for $P_0(t)$

- $P_0(t+\Delta T) = P\{\mathbf{N}(0, t+\Delta T] = 0\}$

  $= P[\{\mathbf{N}(0, t]=0\} \cap \{\mathbf{N}(t, t+\Delta T]=0\}]$

  $= P\{\mathbf{N}(0, t] = 0\} \bullet P\{\mathbf{N}(t, t+\Delta T] = 0\}$

  independence of disjoint intervals

  $= P_0(t) \bullet (1 - \mu \bullet \Delta T)$

  $P\{\mathbf{N}(0, t] = 0\} = P_0(t)$ by definition

  $P\{\mathbf{N}(t, t+\Delta T]\} = 0\} = 1 - \mu \bullet \Delta T$
- $[P_0(t+\Delta T) - P_0(t)]/\Delta T = -\mu \bullet P_0(t)$

Solving the diff. eq. for $P_0(t)$

- $[P_0(t+\Delta T) - P_0(t)]/\Delta T = -\mu \bullet P_0(t)$
- $\dfrac{dP_0(t)}{dt} = -\mu \bullet P_0(t)$
- This is a first-order linear differential equation with constant coefficients
- Initial condition:

  $P_0(0) = P\{$no arrivals in $(0,0]\}$

  $= P\{$no arrivals in $\} = 1$ (not zero!!)
- $P_0(t) = \exp(-\mu \bullet t)$ for $t \ge 0$

Thoughts about $P_0(t)$

- $P_0(t) = P\{$no arrivals in $(0, t]$
  $= P\{\mathbf{N}(0, t] = 0\} = P\{\mathbf{X}_1 > t\}$
  $= \exp(-\mu \bullet t)$
- $P_0(t)$ decays away exponentially as $t \to \infty$
- $P_0(\Delta t) = P\{$no arrivals in $(0, \Delta t] = \exp(-\mu \bullet \Delta t)$
  $= 1 - \mu \bullet \Delta t + (\mu \bullet \Delta t)^2/2! - \ldots$
  $\approx 1 - \mu \bullet \Delta t$ for small values of Δt
- $P\{\mathbf{X}_1 > t\} = \exp(-\mu \bullet t)$ for $t \geq 0$

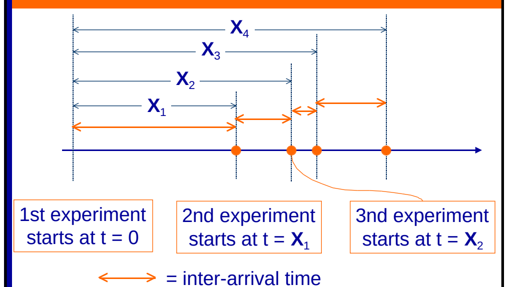
Distribution of the first arrival time

- The first arrival occurs at random time $\mathbf{X}_1$
- $P\{\mathbf{X}_1 > t\} = \exp(-\mu \bullet t)$ for $t \geq 0$
- $P\{\mathbf{X}_1 \leq t\} = F_{\mathbf{X}_1}(t) = 1 - \exp(-\mu \bullet t)$ for $t \geq 0$
- $f_{\mathbf{X}_1}(t)$ = derivative of CDF $F_{\mathbf{X}_1}(t)$
  $= \mu \bullet \exp(-\mu \bullet t)$ for $t \geq 0$
- The first arrival time $\mathbf{X}_1$ is an exponential random variable with parameter μ
- $E[\mathbf{X}_1] = 1/\mu$

Distribution of the inter-arrival time

- The first arrival occurs at time $\mathbf{X}_1 = \tau$
- At $t = \tau$, the experiment begins again
- $\mathbf{X}_2$ = second arrival time
- $\mathbf{X}_2 - \mathbf{X}_1$ is the inter-arrival time
- Given $\mathbf{X}_1 = \tau$, we set up a similar diff. eq. for P{no arrivals in $(\tau, \tau + t]$}
- The inter-arrival time $\mathbf{X}_2 - \mathbf{X}_1$ is an exponential RV with parameter μ

An illustrative diagram



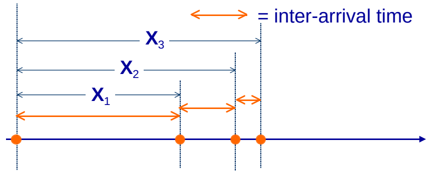
Distribution of inter-arrival times

- Experiment begins again after each arrival
- Successive experiments are repeated independent trials
- $\mathbf{X}_k - \mathbf{X}_{k-1}$ is k-th inter-arrival time
- Proceeding as before, all the inter-arrival times are independent exponential RVs with parameter μ
- Is $\mathbf{X}_1$ also an inter-arrival time?
  Yes, if there was an arrival at $t = 0$

Distribution of waiting times

- $\mathbf{X}_1$ is an inter-arrival time if there was an arrival at $t = 0$
- Otherwise, $\mathbf{X}_1$ is the waiting time for the next arrival after $t = 0$
- pdf of $\mathbf{X}_1$ is same as pdf of inter-arrivals
- Generally, waiting time for the next arrival after $t = \tau$ also has the same pdf
- It is not necessary to have an arrival at τ for this result to hold

An illustrative diagram



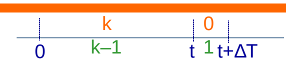
1st experiment starts at t = 0

X_1 is an inter-arrival time if there was an arrival at t = 0; a waiting time otherwise

Distribution of the k-th arrival time

- The k-th arrival occurs at random time $\mathbf{X}_k$
- We set up and solve a differential equation for $P_k(t)$ = P{exactly k arrivals in (0, t]}
  = P{**N**(0, t] = k}
- $P_k(t+\Delta T)$ = P{k arrivals in (0, t+ΔT]
  = P{**N**(0, t+ΔT] = k}
  = P[{**N**(0, t] = k} ∩ {**N**(t, t+ΔT] = 0}]
  + P[{**N**(0, t] = k–1}∩P{**N**(t, t+ΔT] = 1}]

The differential equation for $P_k(t)$



- $P_k(t+\Delta T)$
  = P[{**N**(0, t] = k} ∩ {**N**(t, t+ΔT] = 0}]
  + P[{**N**(0, t] = k–1}∩P{**N**(t, t+ΔT] = 1}]
  $= P_k(t)\bullet(1 - \mu\bullet\Delta T) + P_{k-1}(t)\bullet(\mu\bullet\Delta T)$
  $[P_k(t+\Delta T) - P_k(t)]/\Delta T = -\mu\bullet P_k(t) + \mu\bullet P_{k-1}(t)$
- $\frac{dP_k(t)}{dt} = -\mu\bullet P_k(t) + \mu\bullet P_{k-1}(t)$

Solving the diff. eq. for $P_k(t)$

- $\frac{dP_k(t)}{dt} = -\mu\bullet P_k(t) + \mu\bullet P_{k-1}(t)$
- We can set k = 1 and solve for $P_1(t)$ since we know that $P_0(t) = \exp(-\mu\bullet t)$
- Next, set k = 2 and solve for $P_2(t)$ since we know $P_1(t)$,
- and so on
- Alternatively, we can use LaPlace transforms

LaPlace transforms for $P_k(t)$

- $\frac{dP_k(t)}{dt} = -\mu\bullet P_k(t) + \mu\bullet P_{k-1}(t)$
- Initial condition:
  $P_k(0)$ = P{k arrivals in (0,0]}
  = P{k arrivals in   } = 0 (not one!!)
- $s\bullet£[P_k(t)] = -\mu\bullet£[P_k(t)] + \mu\bullet£[P_{k-1}(t)]$
- $£[P_k(t)] = \frac{\mu}{s+\mu}£[P_{k-1}(t)] = \left(\frac{\mu}{s+\mu}\right)^2£[P_{k-2}(t)]$
  $\ldots = \left(\frac{\mu}{s+\mu}\right)^k\bullet£[P_0(t)] = \left(\frac{\mu}{s+\mu}\right)^k\left(\frac{1}{s+\mu}\right)$

Invert LaPlace transform to get $P_k(t)$

- $£[P_k(t)] = \frac{\mu}{s+\mu}£[P_{k-1}(t)] = \left(\frac{\mu}{s+\mu}\right)^2£[P_{k-2}(t)]$
  $\ldots = \left(\frac{\mu}{s+\mu}\right)^k\bullet£[P_0(t)] = \left(\frac{\mu}{s+\mu}\right)^k\left(\frac{1}{s+\mu}\right)$
- $P_k(t) = \frac{(\mu\bullet t)^k}{k!}\exp(-\mu\bullet t)$ = P{**N**(0, t] = k}
- For any fixed value of t,
  P{**N**(0, t] = k} $= \frac{(\mu\bullet t)^k}{k!}\exp(-\mu\bullet t)$
- **N**(0, t] is a Poisson RV with parameter $\mu\bullet t$

It's a Poisson process!

- For any fixed value of t,
  $P\{\mathbf{N}(0, t] = k\} = \frac{(\mu \cdot t)^k}{k!} \exp(-\mu \cdot t)$
- $\mathbf{N}(0, t]$ is a Poisson RV with parameter $\mu \cdot t$
- Start counting arrivals after time t_1
- Then $\mathbf{N}(t_1, t_2]$ is a Poisson RV with parameter $\mu(t_2 - t_1) = \mu \cdot$duration of interval
- The RVs $\mathbf{N}(t_1, t_2]$, $\mathbf{N}(t_3, t_4]$, $\mathbf{N}(t_5, t_6]$, … are independent if the intervals $(t_1, t_2]$, $(t_3, t_4]$, $(t_5, t_6]$, … don't overlap

Distribution of the k-th arrival time

- The k-th arrival occurs at random time $\mathbf{X}_k$
- $P\{\mathbf{X}_k > t\} = P\{\mathbf{N}(0, t] \le k-1\}$

$= \exp(-\mu t) \cdot [1 + (\mu t) + (\mu t)^2/2! + \ldots + (\mu t)^{k-1}/(k-1)!]$

- $f_{\mathbf{X}_k}(t)$ = derivative of CDF $F_{\mathbf{X}_k}(t)$
  $= -$ derivative of $P\{\mathbf{X}_k > t\}$
  $= \mu \cdot ((\mu t)^{k-1}/(k-1)!) \cdot \exp(-\mu \cdot t)$ for $t > 0$
  $= \mu \cdot \exp(-\mu t) \cdot (\mu t)^{k-1}/\Gamma(k)$
- The k-th arrival time $\mathbf{X}_k$ is a gamma random variable with parameters (k, μ)

Waiting time for next k arrivals

- The k-th arrival time $\mathbf{X}_k$ is a gamma random variable with parameters (k, μ)
- Starting at any time $t = \tau$, the waiting time for the next k arrivals has the same pdf
- It is not necessary to have an arrival at τ for this result to hold
- $E[\mathbf{X}_k] = k/\mu = k \cdot E[\mathbf{X}_1]$
- Average waiting time for k arrivals
  = k•{average waiting time for one arrival}

A final observation…

- The k-th arrival time $\mathbf{X}_k$ is a gamma random variable with parameters (k, μ)
- But, $\mathbf{X}_k = (\mathbf{X}_k - \mathbf{X}_{k-1}) + (\mathbf{X}_{k-1} - \mathbf{X}_{k-2}) + (\mathbf{X}_{k-2} - \mathbf{X}_{k-3}) + \ldots + (\mathbf{X}_2 - \mathbf{X}_1) + \mathbf{X}_1$
  is the sum of k inter-arrival times, i.e. k independent $(1, \mu)$ gamma RV
- Special case of a general result: The sum of independent gamma RVs with same scale parameter is a gamma RV
- order = sum of the orders; scale is same

Summary

- We have discussed the Poisson process in some detail
- Basic assumptions:
  - Arrival rate is μ
  - P{one arrival in ΔT interval} $\approx \mu \cdot \Delta T$
  - P{no arrival in ΔT interval} $\approx 1 - \mu \cdot \Delta T$
  - Arrivals in disjoint time intervals are independent (process is memoryless)

Consequences of assumptions

- Inter-arrival times are exponential random variables with parameter μ
- Average inter-arrival time is $1/\mu$
- Time of k-th arrival/waiting time for next k is a gamma RV with parameters (k, μ)
- Average waiting time for k-th arrival is k/μ
- Number of arrivals in a time interval of length t is a Poisson random variable with parameter $\mu \cdot t$

