What is a Poisson process?

- Consider a sequence of random arrivals (occurrences of an event of interest)
- The time between two successive arrivals varies randomly: it is a random variable
- The inter-arrival times are independent RV arising from independent trials
- Observed average value of the inter-arrival times = {time of the N-th arrival}/N
- Observed average ≈ expected value of RV

Arrival rate μ

- Let $1/\mu$ denote the expected value of the inter-arrival time
- For large N, the total of N inter-arrival times is roughly $N \cdot (1/\mu) = T$
- The number of arrivals in a long time interval of duration T is roughly $N = \mu T$
- μ is called the arrival rate or intensity
- On average, there are μT arrivals in an interval of duration T

Basic assumptions

- At any time instant t, at most one arrival can occur
- Assumptions: For small values of ΔT
  - $P\{\mathbf{N}(t, t+\Delta t] = 1\} = \mu \cdot \Delta T$
  - $P\{\mathbf{N}(t, t+\Delta t] = 0\} = 1 - \mu \cdot \Delta T$
  - The numbers of arrivals in disjoint intervals (i.e. non-overlapping intervals) of time are independent

P{no arrivals in (0, t]}

- The first arrival occurs at random time $\mathbf{X}_1$
- $P_0(t) = P\{\text{no arrivals in } (0, t]\} = P\{\mathbf{X}_1 > t\}$
- $\dfrac{dP_0(t)}{dt} = -\mu \cdot P_0(t)$; $P_0(0) = 1$
- $P_0(t) = \exp(-\mu \cdot t)$ for $t \geq 0$
- $P_0(\Delta t) = P\{\text{no arrivals in } (0,\Delta t] = \exp(-\mu \cdot \Delta t)$
  $\approx 1 - \mu \cdot \Delta t$ for small values of Δt
- $P\{\mathbf{X}_1 > t\} = \exp(-\mu \cdot t)$ for $t \geq 0$

Distribution of the first arrival time

- The first arrival occurs at random time $\mathbf{X}_1$
- $P\{\mathbf{X}_1 > t\} = \exp(-\mu \cdot t)$ for $t \geq 0$
- Complementary CDF of $\mathbf{X}_1$ is the same as that of an exponential RV with parameter μ
- $f_{\mathbf{X}_1}(t) = -$ derivative of complementary CDF
  $= \mu \cdot \exp(-\mu \cdot t)$ for $t \geq 0$
- $\mathbf{X}_1$ is an exponential RV with parameter μ
- $E[\mathbf{X}_1] = 1/\mu$

P{exactly k arrivals in (0, t]}

- The k-th arrival occurs at random time $\mathbf{X}_k$
- $P_k(t) = P\{\text{exactly } k \text{ arrivals in } (0,t]\}$
- $\dfrac{dP_k(t)}{dt} = -\mu \cdot P_k(t) + \mu \cdot P_{k-1}(t)$; $P_k(0) = 0$
- Use LaPlace transforms to solve
- $P_k(t) = \dfrac{(\mu \cdot t)^k}{k!} \exp(-\mu \cdot t) = P\{\mathbf{N}(0, t] = k\}$
  where $\mathbf{N}(0, t]$ = number of arrivals in (0, t] is a discrete random variable
- $\mathbf{N}(0, t]$ is a Poisson RV with parameter $\mu \cdot t$

Distribution of the k-th arrival time

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

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

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

It's a Poisson process!

- For any fixed value of t,
  $P\{\mathbf{N}(0, t] = k\} = \frac{(\mu\bullet t)^k}{k!}\exp(-\mu\bullet t)$
- $\mathbf{N}(0, t]$ is a Poisson RV with parameter $\mu\bullet t$
- $\mathbf{N}(t_1, t_2]$ = number of arrivals in $(t_1, t_2]$ is a Poisson RV with parameter $\mu(t_2-t_1)$
- The RVs $\mathbf{N}(t_1, t_2]$, $\mathbf{N}(t_3, t_4]$, $\mathbf{N}(t_5, t_6]$, … are independent if the corresponding intervals $(t_1, t_2]$, $(t_3, t_4]$, $(t_5, t_6]$, … are disjoint

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
- $E[\mathbf{X}_k] = k/\mu = k\bullet E[\mathbf{X}_1]$
- Average waiting time for k arrivals
  = k•{average waiting time for one arrival}

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

Example

- The stream of packets arriving at a router is modeled as a Poisson process with arrival rate 5 per second
- On average, how many packets arrive at the router in one minute?
- Average = arrival rate•time = 5•60 = 300
- P{12 packets in the next second}?
- # packets is Poisson RV with parameter 5
- P{12 packets in second} = $\exp(-5)\bullet 5^{12}/12!$

An interesting result – set up

- Given that M packets arrived in (0, t], what is the conditional probability that the first k packets arrived in $(0, \tau]$, where $\tau < t$?
- P{k packets in $(0, \tau]$ | M packets in (0, t]}

$$= \frac{P(\{\mathbf{N}(0, \tau] = k\} \cap \{\mathbf{N}(0, t] = M\})}{P\{\mathbf{N}(0, t] = M\}}$$

$$= \frac{P(\{\mathbf{N}(0, \tau] = k\} \cap \{\mathbf{N}(\tau, t] = M-k\})}{P\{\mathbf{N}(0, t] = M\}}$$

$$= \frac{P\{\mathbf{N}(0, \tau] = k\}\bullet P\{\mathbf{N}(\tau, t] = M-k\}}{P\{\mathbf{N}(0, t] = M\}}$$

An interesting result — nearly done

- For $\tau < t$, P{**N**(0, τ] = k | **N**(0, t] = M}
  = P{k packets in (0, τ] | M packets in (0, t]}
  $= \dfrac{P\{\mathbf{N}(0,\tau] = k\}\bullet P\{\mathbf{N}(\tau, t] = M - k\}}{P\{\mathbf{N}(0, t] = M\}}$
- **N**(0, τ], **N**(τ, t] and **N**(0, t] are Poisson RVs with parameters μτ, μ(t–τ), μt respectively
- P{k packets in (0, τ] | M packets in (0, t]} =

$$\frac{(\mu\tau)^k \exp(-\mu\tau)/k! \bullet (\mu(t-\tau))^{M-k}\exp(-\mu(t-\tau))/(M-k)!}{(\mu t)^M \exp(-\mu t)/M!}$$

An interesting result — finally!

- P{k packets in (0, τ] | M packets in (0, t]} =

$$\frac{(\mu\tau)^k \exp(-\mu\tau)/k! \bullet (\mu(t-\tau))^{M-k}\exp(-\mu(t-\tau))/(M-k)!}{(\mu t)^M \exp(-\mu t)/M!}$$

$$= \frac{M!\tau^k \bullet (t-\tau)^{M-k}}{k! \bullet (M-k)! \bullet t^M} = \binom{M}{k}\left(\frac{\tau}{t}\right)^k \left(\frac{t-\tau}{t}\right)^{M-k}$$

- This is a binomial (M, τ/t) probability!
- Given M arrivals in interval of length T, the number of packets in subinterval of length τ is a binomial RV with parameters (M, τ/T) no matter what the arrival rate!

Saying it again more generally

- In a Poisson process with arrival rate μ, the number of arrivals in an interval A is a Poisson RV with parameter μ•(length of A)
- However, given that M arrivals occurred in B, where A B, the conditional pmf of the number of arrivals in A is a binomial pmf with parameters (M, length(A)/length(B)) regardless of the value of μ

...and now for something completely different

- We turn our attention from Poisson processes to the most famous of all continuous random variables, the

Gaussian

or

normal

random variable

Definition of Gaussian RV

- The continuous random variable **X** is called a Gaussian or normal random variable with mean μ and variance σ^2 if its pdf is given by
- $f(u) = \dfrac{1}{\sigma\sqrt{2\pi}} \bullet \exp\left(-\dfrac{1}{2}\left(\dfrac{u-\mu}{\sigma}\right)^2\right), \quad -\infty < u < \infty$
- We often write $\mathbf{X} \sim N(\mu, \sigma^2)$ or $\mathbf{X} \sim \mathcal{N}(\mu, \sigma^2)$ or use the phrase "**X** is $N(\mu, \sigma^2)$" to say that **X** is Gaussian: mean μ, variance σ^2

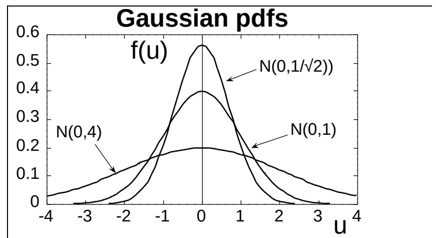
Some thoughts re Gaussian pdfs

- $f(u) = \dfrac{1}{\sigma\sqrt{2\pi}} \bullet \exp\left(-\dfrac{1}{2}\left(\dfrac{u-\mu}{\sigma}\right)^2\right), \quad -\infty < u < \infty$
- $\mathbf{X} \sim N(\mu, \sigma^2)$ or $\mathbf{X} \sim \mathcal{N}(\mu, \sigma^2)$
- The mean and the variance specify the pdf completely
- The pdf is symmetric about u = μ, and has value $1/(\sigma \bullet (2\pi)^{1/2})$ at u = 0
- Is the area under the curve = 1? Yes, read the proof for yourself in Ross, Chapter 5

Unit or standard Gaussian RV

- If $X \sim N(0, 1)$, it is called a unit or standard Gaussian random variable. Its pdf is $\phi(u)$
- $\phi(u) = (2\pi)^{-1/2} \bullet \exp(-u^2/2)$, $-\infty < u < \infty$
- We use $\phi(u)$ to denote the unit Gaussian pdf and $\Phi(u)$ to denote the corresponding unit Gaussian CDF
- $f(u) = \frac{1}{\sigma\sqrt{2\pi}} \bullet \exp\left(-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2\right)$, $-\infty < u < \infty$
  $= \sigma^{-1} \bullet \phi((u-\mu)/\sigma)$

Gaussian pdfs look like this



"The familiar bell-shaped curve"

Why call it "Gaussian"?

- The Gaussian pdf is named after
  Karl Friedrich Gauss (1777-1855)
  one of the greatest mathematicians of all time
- Gauss used this pdf to study errors in measurement, least-squares curve-fitting, LMSE prediction, etc,: topics that are just as central to modern scientific analysis as they were to 19th century mathematics

Why call it "normal"?

- Many sets of data gathered from a variety of physical phenomena seem to fit the Gaussian (or normal) distribution
- During the past 150 years, this empirical observation has been studied extensively
- Henri Poincaré: Everybody believes in the normal law: the physicists because they think it is a mathematical theorem, and the mathematicians because they think it is an experimentally observed fact

Why are most data sets "normal"?

- The central limit theorem asserts that if many "small" "random" causes produce a net effect, then that effect can be modeled as a normal random variable
- Noise voltages in electrical circuits have normal distributions
- The noise voltage is the net result of the electric fields created by gazillions of electrons in random positions: each charge has a small effect on the voltage

Isn't everything "normal" anyway?

- Lots of people seem to think that all data sets must be "normally" distributed

  and go through lots of contortions to make the data conform
- Example: Raw SAT and GRE scores are "curved" to get a $N(500, 100^2)$ distribution
- A nonlinear transformation is applied to raw scores to ensure that the histogram of the result looks like a $N(500, 100^2)$ pdf

Why not call it normal?

- Calling a Gaussian RV a normal RV is pejorative in the sense that many people take it to mean that all other RVs and pdfs are, somehow, abnormal
- Lots of non-Gaussian distributions are very interesting, and are very appropriate models for some physical phenomena
- We shall use the name Gaussian exclusively

Unit Gaussian CDF Φ(x)

- The unit Gaussian CDF Φ(x) is the area under the unit Gaussian pdf

  $\varphi(u) = (2\pi)^{-1/2} \cdot \exp(-u^2/2), -\infty < u < \infty$

  to the left of the point x
- Φ(x) cannot be expressed in terms of simple functions (e.g. exp, ln, sin) of x
- The antiderivative of $\exp(-u^2/2)$ is NOT $(1/u)\cdot\exp(-u^2/2)$ or $(-1/u)\cdot\exp(-u^2/2)$ or $(u^3/6)\cdot\exp(-u^2/2)$ or …

So, how do we find Φ(x)?

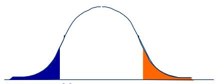
- The values of Φ(x) have been calculated via numerical integration, and are available in tables (see, e.g. Ross, p. 208)
- Φ(x) is a built-in pre-programmed function in many advanced scientific calculators
- Does your calculator know about Φ(x)?
- Do calculators use numerical integration?
- No, a rational function approximation to Φ(x) is used to calculate its value

Rational approximation to Φ(x)

- For $x \geq 0$, Φ(x) $\approx 1 - \varphi(x)\cdot[b_1\cdot t+b_2\cdot t^2+b_3\cdot t^3+b_4\cdot t^4+b_5\cdot t^5]$ where $t = [1 + 0.2316419\cdot x]^{-1}$ and
  - $b_1 = 0.319381530$
  - $b_2 = -0.356563782$
  - $b_3 = 1.781477937$
  - $b_4 = -1.821255978$
  - $b_5 = 1.330274429$
- Approximation error magnitude $\leq 7.5\cdot 10^{-8}$

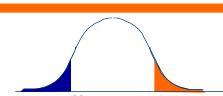
Properties of $\Phi(x)$

- $\Phi(x)$ increases from 0 at $-\infty$ to 1 at $+\infty$
- $\Phi(0) = 1/2$
- $\Phi(x)$ is the area to the left of x under the even (symmetric about 0) function $\varphi(u)$



- $1-\Phi(x)$ = orange area = $\Phi(-x)$ = blue area

Properties of $\Phi(x)$ (continued)



- $1-\Phi(x)$ = orange area = $\Phi(-x)$ = blue area
- $\Phi(-x) = 1 - \Phi(x)$
- $\Phi(x)$ is tabulated only for $x \geq 0$
- $\Phi(3) = 0.9987$ • $\Phi(1.96) - \Phi(-1.96) = 0.95$
- Most of the probability mass lies in the interval $[-3, 3]$

$Q(x) = 1 - \Phi(x)$

- $Q(x) = 1 - \Phi(x)$ = area to the right of x under $\varphi(u)$ is the complementary unit Gaussian CDF function
- In many applications, answers are given in terms of $Q(\bullet)$ rather than $\Phi(\bullet)$
- Example: The bit error probability for coherent binary orthogonal frequency-shift-keying data transmission is $Q(\sqrt{SNR})$ where SNR is the signal-to-noise ratio

More about $Q(x) = 1 - \Phi(x)$

- $Q(x) = 1 - \Phi(x)$ decreases from 1 at $-\infty$ to 0 at $+\infty$
- $Q(0) = 1/2$
- For $x > 0$,
  $$(x^{-1} - x^{-3})\bullet\varphi(x) < Q(x) < x^{-1}\bullet\varphi(x)$$
- The upper and lower bounds on $Q(x)$ diverge to $\pm\infty$ as x approaches 0
- $Q(x) \leq (1/2)\bullet\exp(-x^2/2)$ for $x \geq 0$ is not as tight for large x but more useful for small x

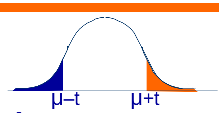
What about non-unit Gaussians?

- If **X** is $N(\mu, \sigma^2)$, its pdf is given by
- $$f(u) = \frac{1}{\sigma\sqrt{2\pi}}\bullet\exp\left(-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2\right), \quad -\infty < u < \infty$$
  $$= \sigma^{-1}\bullet\varphi((u-\mu)/\sigma)$$
  where $\varphi(\bullet)$ is the unit Gaussian pdf
- Claim: $F(u)$ = CDF of **X** $= \Phi((u-\mu)/\sigma)$ where $\Phi(\bullet)$ is the unit Gaussian CDF
- Proof: Use chain rule to find derivative of $\Phi((u-\mu)/\sigma)$. Answer: $\sigma^{-1}\bullet\varphi((u-\mu)/\sigma) = f(u)$

Gaussian CDFs in general

- CDF $F(x)$ of a $N(\mu, \sigma^2)$ RV can be stated in terms of the unit Gaussian CDF $\Phi(x)$
- All probabilities can be expressed in terms of $\Phi(x)$, and obtained from the tables or by using calculators
- Older calculators compute $\Phi(x)$ given x
- Newer calculators compute $F(x)$ given μ, σ, and x (or given μ, σ^2, and x)
- Learn your calculator's preference

An example



- **X** ~ **N**(μ, σ^2) with CDF F(x)

P{**X** > t} = 1 – F(t) = 1 – Φ((t–μ)/σ)

P{**X** < –t} = F(–t) = Φ((–t–μ)/σ) = 1–Φ((t+μ)/σ)

P{|**X**–μ| > t} = P{**X** > μ+t} + P{**X** < μ–t}

= 1 – Φ(t/σ) + Φ(–t/σ)

= 2•[1 – Φ(t/σ)]

Some numerology

- **X** ~ **N**(μ, σ^2) with CDF F(x)
- 68.26% of the probability mass is in the range (μ–σ, μ+σ)
- 95% of the probability mass is in the range (μ–1.96σ, μ+1.96σ)
- 95.44% of the probability mass is in the range (μ–2σ, μ+2σ)
- 99.74% of the probability mass is in the range (μ–3σ, μ+3σ)

The DeMoivre-LaPlace theorem

- Let **X** denote a binomial random variable with parameters (n, p)
- E[**X**] = np = μ
- var(**X**) = np(1–p) = σ^2
- For large values of n,
- P{a < **X** < b} ≈ Φ((b–μ)/σ) – Φ((a–μ)/σ)
- Approximation is good if μ (a ,b) but can be bad otherwise
- For example, approximation assigns nonzero probability to {**X** < 0} and {**X** > n}

DeMoivre-LaPlace theorem

- The DeMoivre-LaPlace theorem is a special case of the central limit theorem
- The theorem can be used to reduce the lengths of confidence intervals for a given confidence level
- Actually, very many fewer people need to be polled than we previously obtained via the Chebyshev inequality

Summary

- We took a brief second look at Poisson processes and discovered an interesting result about a conditional pmf
- We discussed Gaussian (a.k.a. normal) random variables
- All Gaussian CDFS can be expressed in terms of the unit Gaussian CDF
- Values of Gaussian CDFs must be computed; the values cannot be obtained analytically in terms of simpler functions