Continuous Probability Distributions

Exponential, Erlang, Gamma

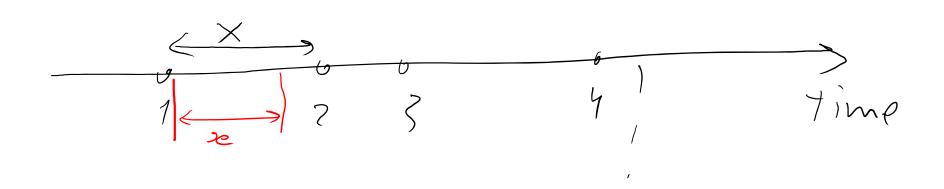
Poisson process discrete events happen at rate) Expected number of events 1h time 20 is la The actual number of events Nz is a Poisson distributed discrete random variable $P(N=n)=(\frac{\lambda x}{1})^n e^{-\lambda x}$ Why Poisson? Divide X into many tiny intervals of Length DX Prob(N=n)=(L)pn(1-p)Lin $p = \lambda s x$ l = x/s xU p~orc→o, L~ bx→∞ $E(N_{E}) = \rho L = \lambda x$ Poisson

Poisson (constant rate) processes

- Let's assume that proteins are produced by all ribosomes in the cell at a rate λ per second.
- The expected number of proteins produced in x seconds is λx .
- The actual number of proteins N_x is a discrete random variable following a Poisson distribution with mean λx :

$$P_N(N_x=n)=\exp(-\lambda x)(\lambda x)^n/n!$$
 $E(N_x)=\lambda x$

- Why Discrete Poisson Distribution?
 - Divide time into many tiny intervals of length $x_0 << 1/\lambda$
 - The probability of success (protein production) per internal is small: $p = \lambda \cdot \Delta x << 1$,
 - The number of intervals is large: L= $x/\Delta x >> 1$
 - Mean is constant: $E(N_x)=p \cdot L=(\lambda \Delta x) \cdot (x/\Delta x)=\lambda \cdot x$
 - $P(N_x=n)=L!/n!(L-n)! p^n (1-p)^{L-n}$
 - − In the limit $p \rightarrow 0$, $L \rightarrow \infty$: Binomial distribution \rightarrow Poisson



$$CCTF; \Pr(X>x) = \Pr(N_x = 0) = \frac{1}{\sqrt{2}} e^{-\lambda x} = e^{-\lambda x}$$

$$= \frac{\sqrt{2}}{\sqrt{2}} e^{-\lambda x} = e^{-\lambda x}$$

$$PDF = -\frac{1}{\sqrt{2}} CTF; \qquad f(x) = \lambda e^{-\lambda x}$$

What is the distribution of the interval X between CONSEQUITIVE EVENTS of a constant rate process?

- X is a continuous random variable
- CCDF: $Prob(X>x) = Prob(N_X=0) = exp(-\lambda x)$.
 - Remember: $P_N(N_x=n)=exp(-\lambda x) (\lambda x)^n/n!$
- PDF: $f(x) = -d \ CCDF(x)/dx = \lambda exp(-\lambda x)$
- We started with a discrete Poisson distribution where time x was a parameter and N_x – discrete random variable
- We ended up with a continuous exponential distribution where time X between events was a continuous random variable

Exponential Mean & Variance

If the random variable X has an exponential distribution with parameter λ ,

$$\mu = E(X) = \frac{1}{\lambda}$$
 and $\sigma^2 = V(X) = \frac{1}{\lambda^2}$ (4-15)

Note that, for the:

- Poisson distribution: mean= variance
- Exponential distribution: mean = standard deviation = variance^{0.5}

Exponential Distribution is a continuous generalization of what discrete probability distribution?

- A. Poisson
- B. Binomial
- C. Geometric
- D. Negative Binomial
- E. I have no idea

Get your i-clickers

Biochemical Reaction Time

 The time x (in minutes) until an enzyme successfully catalyzes a biochemical reaction is approximated by this CDF:

$$F(x) = 1 - e^{-x/1.4}$$
 for $0 \le x$

What is the PDF?

$$f(x) = \frac{dF(x)}{dx} = \frac{d}{dx} [1 - e^{-x/1.4}] = e^{-x/1.4}/1.4 \text{ for } 0 \le x$$

What proportion of reactions is complete within 0.5 minutes?

$$P(X < 0.5) = F(0.5) = 1 - e^{-0.5/1.4} = 1 - 0.7 = 0.3$$

The reaction product is "overdue": no product has been generated in the past 3 minutes.

What is the probability that a product will appear in the next 0.5 minutes?

- B. 0.3
- C. 0.62
- D. 0.99
- E. I have no idea

$$F(x) = 1 - e^{-x/1.4}$$

$$F(0.5) \approx 0.3$$

$$F(3.5) \approx 0.92$$

Get your i-clickers

Memoryless property of the exponential P(X>t+s|X>s) = P(X>t) $P(X>t+s \mid X>s) = \frac{P(X>t+s, X>s)}{P(X>s)} =$ $=\frac{e\times p(-\lambda(t+s))}{e\times p(-\lambda s)}=e\times p(-\lambda t)=$ $= \mathcal{P}(X > t)$ Exponential is the only memoryless distribution

Can other random variables be nemory less? $\mathcal{P}(X>S+t|X>S) = \mathcal{P}(X>t)$ $\frac{P(X>S+t)}{P(X>S)} = P(X>t)$ P(X>S++) = P(X>S).P(X>t)for any + & S Let t= as - very small; +(s)=P(X>s) F(S+DS)= F(DS). F(S) F(0)=1=> = (05)=1-) as

$$F(S+\Delta S) = (1-\lambda \Delta S) F(S)$$

$$F(S+\Delta S) - F(S)$$

$$\Delta S = -\lambda F(S)$$

$$\frac{dF(S)}{dS} = -\lambda F(S)$$

$$F(S) = \exp(-\lambda S)$$

$$PDF(S) = -\frac{dF}{dS} = \lambda \exp(-\lambda S)$$
Thus, any continuous reverses r.v.
$$Discrete = geometric$$

Exponential Distribution in Reliability

- The reliability of electronic components is often modeled by the exponential distribution. A chip might have mean time to failure of 40,000 operating hours.
- The memoryless property implies that the component does not wear out – the probability of failure in the next hour is constant, regardless of the component age.
- The reliability of mechanical components do have a memory – the probability of failure in the next hour increases as the component ages.

Erlang Distribution

- The Erlang distribution is a generalization of the exponential distribution.
- The exponential distribution models the time interval to the 1st event, while the
- Erlang distribution models the time interval to the rth event, i.e., a sum of r exponentially distributed variables.
- The exponential, as well as Erlang distributions, is based on the constant rate Poisson process.

Erlang Distribution

Generalizing from the constant rate Poisson → Exponential:

$$P(X > x) = \sum_{k=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!} = 1 - F(x)$$

Now differentiating F(x) we find that all terms in the sum except the last one cancel each other:

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{(r-1)!} \quad \text{for } x > 0 \quad \text{and} \quad r = 1, 2, 3, \dots$$

$$\sum_{\text{Sec } 4-9 \text{ Erlang & Gamma Distributions}} r = 3$$

Example 4-23: Medical Device Failure

The failures of medical devices can be modeled as a Poisson process. Assume that units that fail are repaired immediately and the mean number of failures per hour is 0.0001. Let X denote the time until 4 failures occur. What is the probability that X exceed 40,000 hours ~=4.5 years?

Let the random variable N denote the number of failures in 40,000 hours. The time until 4 failures occur exceeds 40,000 hours iff the number of failures in 40,000 hours is ≤ 3 .

$$P(X > 40,000) = P(N \le 3)$$

 $E(N) = 40,000(0.0001) = 4$ failures in 40,000 hours
 $P(N \le 3) = \sum_{k=0}^{3} \frac{e^{-4}4^k}{k!} = 0.433$

Erlang Distribution is a continuous generalization of what discrete probability distribution?

- A. Poisson
- B. Binomial
- C. Geometric
- D. Negative Binomial
- E. I have no idea

Get your i-clickers

 $F(x) = \frac{\chi^r \chi^{r-1} \exp(-\chi x)}{(r-1)!}$ Can be generalized for any r > 0

Q: What to use instead of (r-1)!

Gamma Function

The gamma function is the generalization of the factorial function for r > 0, not just non-negative integers.

$$\Gamma(r) = \int_{0}^{\infty} x^{r-1} e^{-x} dx$$
, for $r > 0$ (4-17)

Properties of the gamma function

$$\Gamma(r) = (r-1)\Gamma(r-1)$$
 recursive property

$$\Gamma(r) = (r-1)!$$
 factorial function

$$\Gamma(1) = 0! = 1$$

$$\Gamma(1/2) = \pi^{1/2} = 1.77$$

$$\Gamma(3/2) = \frac{1}{7} \Gamma(\frac{1}{2}) = 0.39$$
interesting facts
$$\Gamma(3/2) = \frac{1}{7} \Gamma(\frac{1}{2}) = 0.39$$
Sec 4-9 Erlang & Gamma Distributions
$$\Gamma(3/2) = \frac{1}{7} \Gamma(3/2) = 0.39$$

Gamma Distribution

The random variable *X* with a probability density function:

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)}, \text{ for } x > 0$$
 (4-18)

has a gamma random distribution with parameters $\lambda > 0$ and r > 0. If r is an positive integer, then X has an Erlang distribution.

Gamma Distribution Graphs

- The r and λ parameters are often called the "shape" and "scale"
- Different parameter combinations change the distribution.
- The distribution becomes progressively more symmetric as *r* increases.
- Matlab uses 1/ λ as a "scale" parameter.

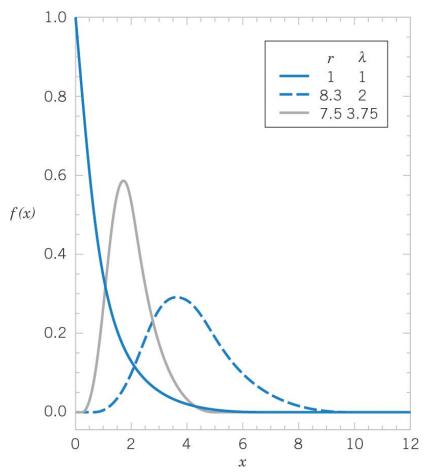


Figure 4-25 Gamma probability density functions for selected values of λ and r.

Mean & Variance of the Erlang and Gamma

 If X is an Erlang (or more generally Gamma) random variable with parameters λ and r,

$$\mu = E(X) = r / \lambda$$
 and $\sigma^2 = V(X) = r / \lambda^2$ (4-19)

• Generalization of exponential results:

$$\mu = E(X) = 1 / \lambda$$
 and $\sigma^2 = V(X) = 1 / \lambda^2$ or Negative binomial results:

$$\mu = E(X) = r / p$$
 and $\sigma^2 = V(X) = r(1-p) / p^2$

Matlab exercise:

- Generate a sample of 100,000 random numbers drawn from an exponential distribution with rate lambda=0.1. Hint: read the help page for random('Exponential'...)
- Calculate mean and standard deviation of the sample and compare to predictions 1/lambda
- Generate PDF and CCDF of the sample and plot them both on a semilogarithmic scale (y-axis)
- After done with exponential modify for Gamma distribution with lambda=0.1, r=4.5

```
Stats=??; lambda=??;

    r2=random('Exponential', ??, Stats,1);

disp([mean(r2),??]);
disp([std(r2),??]);
• %%
step=0.1; [a,b]=hist(r2,0:step:max(r2));

    pdf_e=a./sum(a).?? step;

figure; subplot(1,2,1); semilogy(b,pdf_e,'ko-');
• %%
X=0:0.01:100;
for m=1:length(X);
    ccdf_e(m)=sum(r2 ?? X(m))./Stats;
end;
subplot(1,2,2); semilogy(X,ccdf_e,'ko-');
```

WHY DO WHALES JUMP & WHY ARE WITCHES GREEN WHY ARE THERE MIRRORS ABOVE BEDS WHY IS SEA SALL DELITED & SUMY ARE THERE TREES IN THE MIDDLE OF FIELDS & WHY IS THERE NOT A POKEMON MMO TO SERVE I AUGHING IN TV SHOWS WHY ARE THERE DOORS ON THE FREEWAY # 18 WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA WHY ARE THERE SCARY SOUNDS IN MINECRAFT WHY ISTHERE KICKING IN MY STOMACH WHY ARE THERE TWO SLASHES AFTER HTTP WHY ARE THERE CELEBRITIES, DO OYSTERS HAVE PEARLS WHY DO THEY CALL IT THE CLAP WHY ARE THE AVENGERS FIGHTING THE X MEN 5 WHY ARE KYLE AND CARTMAN FRIENDS WHY IS WOLVERINE NOT IN THE AVENGERS \$

WHY IS THERE AN ARROW ON AANG'S HEAD WHY ARE TEXT MESSAGES BLUE WHY ARE THERE MUSTACHES ON CLOTHES (

Credit: XKCD comics

WHY ARE THERE SLAVES IN

WHY DO TWINS HAVE DIFFERENT FINGERPRINTS & WHY IS HTTPS OROSSED OUT IN RED WHY ARE AMERICANS AFRAID OF DRAGONS WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK

SWHY ARE THERE SURPRIS OF CHATES AND SWHEET PHILEGHOUS L

WHY ARE THERE

GHOSTS

₹WHY IS HTTPS IMPORTANT ONALL

WHY AREN'T MY ARMS GROWING

WHY ARE THERE SO MANY CROWS IN ROCHESTER, MIN

WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY ARE THERE MUSTACHES ON CARS I WHY IS EARTH TILTED & WHY ARE THERE MUSTACHES EVERYWHERE

WHY ARE THERE BRIDESMAIDS WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO DYING PEOPLE REACH UP WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY AREN'T THERE MARGOSE ARTERIES TO A MY HOUSE
WHY AREN'T THERE MARGOSE ARTERIES TO A MY HOUSE マWHY DO SPIDERS CON IS WHY ARE THERE HUGE SPIDERS IN MY HOUSE WHY ARE THERE

뉜 WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE 包WHY ARE THERE SPIDERS IN MY ROOM AWHY ARE THERE SO MANY SPIDERS IN MY ROOM

DYING 50

 $\overline{m{\eta}}$ Why is there no GPS in Laptops $m{arepsilon}$ OWHY DO KNEES CLICK 子 WHY IS PROGRAMMING SO HARD WHY AREN'T THERE E. GRADES TO WHY IS THERE A O OHN RESISTER WHY AREN'T THERE E. GRADES TO WHY DO ANGERICANS HATE SOCCER WHY IS ISOLATION BAD WHY DO RHYMES SOUND GOOD WHY DO BOY'S LIKE ME WHY DON'T BOY'S LIKE ME WHY IS THERE NO SOUND ON CAN WHY IS THERE ALWAYS A JAVA UPDATE TO WHY AREN'T BULLETS SHARP WHY ARE THERE RED DOTS ON MY THIGHS WHY AREN'T BULLETS SHARP WHY IS LYING GOOD THE

WHY IS SEX **50 IMPORTANT** WHY IS THERE AN OWL ON THE DOLLAR BILL WHY ARE THERE TWO SPOCKS

YS WET S

WHY AREN'T MY QUAIL LAYING EGGS WHY ARE ULTRASOUNDS IMPORTANT WHY AREN'T MY QUAIL EGGS HATCHING WHY IS STEALING WRONG {idwhy aren't there any foreign military bases in america

WHY ARE CIGARETTES LEGAL WHY ARE THERE DUCKS IN MY POOL WHY IS JESUS WHITE G WHY DO Q TIPS FEEL GOOD Z



SQUIRRELS

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE WHY IS THERE LIQUID IN MY EAR

> WHY AREN'T THERE GUNS IN HARRY POTTER