# There will be no lecture on 10/15/2024

# The next lecture is next Thursday 10/17/2024

## **Erlang Distribution**

**Generalizes the Exponential Distribution:** waiting time between event 0 and event k (constant rate process with rate=r)

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

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

$$
f(x) = \frac{r^k x^{k-1} e^{-rx}}{(k-1)!}
$$
 for  $x > 0$  and  $k = 1, 2, 3, ...$ 

# Gamma Distribution

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

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

$$
(4-18)
$$

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



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

$$
\int_{0}^{+\infty} f(x) dx = 1, \quad \text{Hence}
$$

$$
\Gamma(k) = \int_{0}^{+\infty} r^{k} x^{k-1} e^{-rx} dx = \int_{0}^{+\infty} y^{k-1} e^{-y} dy
$$

Comparing with Erlang distribution for integer k one gets  $\Gamma(k) = (k-1)!$ 

#### **Gamma Function**

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

$$
\Gamma(k) = \int_{0}^{\infty} y^{k-1} e^{-y} dy, \qquad \text{for } r > 0 \qquad (4-17)
$$

Properties of the gamma function

$$
\Gamma(1) = 1
$$
  
\n
$$
\Gamma(k) = (k - 1)\Gamma(k - 1)
$$
 recursive property  
\n
$$
\Gamma(k) = (k - 1)!
$$
 factorial function  
\n
$$
\Gamma\left(\frac{1}{2}\right) = \pi^{\frac{1}{2}} = 1.77 = \left(-\frac{1}{2}\right)!
$$
 interesting fact





#### $Bernou11$ trials **BERNOULLI FAMILY**



# Mean & Variance of the Erlang and Gamma

• If X is an Erlang (or more generally Gamma) distributed random variable with parameters **r** and *k*,

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

• Generalization of exponential results:  $μ = E(X) = 1/r$  and σ 2 $2 = V(X) = 1/r^2$  or Negative binomial results:  $μ = E(X) = k/p$  and σ 2 = *V* ( *X*) = *k(1-p)* / *p* 2

#### Gamma RV: protein concentrations



Friedman N, Cai L, Xie XS. Phys Rev Lett. 2006;97: 168302.

#### Matlab exercise:

- Generate a sample of 100,000 variables with "Harry Potter" Gamma distribution with  $r = 0.1$  and k=9 $\frac{3}{4}$  (9.75)
- • Calculate mean and compare it to k/r (Gamma)
- Calculate standard deviation and compare it to sqrt(k)/r (Gamma)
- Plot semilog-y plots of PDFs **and CCDFs .**
- Hint: read the help page (better yet documentation webpage) for random('Gamma'…): one of their parameters is different than r

#### Matlab exercise: Gamma

- **Stats=100000; r=0.1; k=9.75;**
- **r2=random('Gamma', k,1./r, Stats,1);**
- **disp([mean(r2),k./r]);**
- **disp([std(r2),sqrt(k)./r]);**
- **step=0.1; [a,b]=hist(r2,0:step:max(r2));**
- **pdf\_g=a./sum(a)./step;**
- **figure;**
- **subplot(1,2,1); semilogy(b,pdf\_g,'ko-'); hold on;**
- **x=0:0.01:max(r2); clear cdf\_g;**
- **for m=1:length(x);**
- •**cdf\_g(m)=sum(r2>x(m))./Stats;**
- •**end;**
- •**subplot(1,2,2); semilogy(x,cdf\_g,'rd-');**

# Continuous Probability Distributions

# Normal or GaussianDistribution



#### Normal or Gaussian Distribution

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

 $-\infty < x < \infty$ 

is a normal random variable

with mean  $\mu$ ,

and standard dewviation  $\sigma$ 

sometimes denoted as



Carl Friedrich Gauss (1777 –1855) German mathematician

 $N(\mu, \sigma)$ 

# Normal Distribution

• The location and spread of the normal are independently determined by mean (μ) and standard deviation (σ)



 $\big( x \big)$ 

1

2

 $\pi\sigma$ 

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

Figure 4-10 Normal probability density functions

Sec 4-6 Normal Distribution 15

 $\left(x-\mu\right)^2$ 

 $\mu$ 

 $2\sigma^2$ 

*x*

 $-11$ 

Gaussian (Normal) distribution is very important because any sum of many independent random variables can be approximated with a Gaussian

# Standard Normal Distribution

• A normal (Gaussian) random variable with

 $\mu$  = 0 and σ  $2 = 1$ 

is called a standard normal random variable and is denoted as *Z*.

• Thed cumulative distribution function of a standard normal random variable is denoted as: $\Phi(z) = P(Z \leq z)$ 

• Values are found in Appendix A Table III to Montgomery and Runger textbook

#### Standardizing

If X is a normal random variable with  $E(X) = \mu$  and  $V(X) = \sigma^2$ , the random variable

$$
Z = \frac{X - \mu}{\sigma} \tag{4-10}
$$

is a normal random variable with  $E(Z) = 0$  and  $V(Z) = 1$ . That is, Z is a standard normal random variable.

and variance  $\sigma^2$ . Suppose  $X$  is a normal random variable with mean  $\mu$ 

Then, 
$$
P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = P(Z \le z)
$$
 (4-11)

where *Z* is a standard normal random variable, and

 $\frac{(x-\mu)}{x}$  is the z-value obta  $z = \frac{X}{X}$  $=\frac{(x-\mu)}{x}$  is the z-value obtained by standardizing x.  $\sigma$ The probability is obtained by using Appendi x Table III

 $P(X < \mu - \sigma) = P(X > \mu + \sigma) = (1 - 0.68)/2 = 0.16 = 16\%$ *P*(*X* < μ - 2σ) =*P*(*X* > μ + 2σ) =(1 - 0.95)/2=0.023=2.3% *P*(*X* < μ - 3σ) =*P*(*X* > μ + 3σ) =(1- 0.997)/2=0.0013=0.13%



Figure 4-12 Probabilities associated with a normal distribution – well worth remembering to quickly estimate probabilities.



#### Standard Normal Distribution Tables

#### Assume *Z* is a standard normal random variable. Find *P* ( *Z* <sup>≤</sup> 1.50). Answer: 0.93319



Find *P* ( *Z* <sup>≤</sup> 0.02). Answer: 0.50398

Sec 4-6 Normal Distribution 21

#### Standard Normal Exercises

- 1. P(Z > 1.26) =1- P(Z < 1.26) =1-0.8962 = =0.1038
- 2.  $P(Z < -0.86) = P(Z > 0.86) = 1 P(Z < 0.86) =$ 1-0.815 =0.195
- 3.  $P(Z > -1.37) = P(Z < 1.37) = 0.915$
- 4.  $P(-1.25 < Z < 0.37) = P(Z < 0.37) P(Z < -1.25)$  $= P(Z < 0.37) - P(Z > 1.25) = P(Z < 0.37) (1-P(Z<1.25))= 0.6443-(1-0.8944)=0.5387$







# Matlab exercise: plot PDF of the Gaussian distribution with **mu=3; sigma=2**

calculate mean, standard deviation and variance,

#### Linear-y and Semilog-y plots of PDF Hint:

Generate Standard normal distribution using randn(Stats,1) then multiply and add using sigma, mu

#### Matlab exercise solution

- **Stats=100000;**
- **mu=3; sigma=2;**
- **r1=sigma.\*randn(Stats,1)+mu;**
- **step=0.1;**
- **[a,b]=hist(r1,(mu-10.\*sigma):step:(mu+10.\*sigma));**
- •**pdf\_n=a./sum(a)./step;**
- **figure; subplot(1,2,1); plot(b,pdf\_n,'ko-');**
- **subplot(1,2,2); semilogy(b,pdf\_n,'ko-');**





## Business buzzword: Six Sigma



### Business literature defined six sigma as no more than 3.4 defective products per million

# Fact checking Six Sigma

- $\bullet$ 1-normcdf(z) computes CCDF in Matlab
- •• I expected Prob(X-μ>6σ)=3.4 defective products per million
- $\bullet$ •  $Prob(X-\mu>6\sigma)=1$ -normcdf(6) = 9.8659e-10 ~1 per billion
- What Six Sigma should be called? Find z such that Prob(X-μ>z · σ)=3.4 per million
- $\bullet$ 5 is not enough: 1-normcdf(5) = 2.8665e-07
- $\bullet$ • 4 is too much: 1-normcd $f(4)$  = 3.1671e-05
- •4.5 is perfect: 1-normcdf(4.5) =3.3977e-06
- Should be called Four-point-Five Sigma but not as catchy

# What's wrong with Six Sigma?

- "Motorola has determined, through years of process and data collection, that processes vary and drift over time – what they call the Long-Term Dynamic Mean Variation. This variation typically falls between 1.4 and 1.6." They shifted their sigma down by 1.5.
- The statistician Donald J. Wheeler has dismissed the 1.5 sigma shift as "goofy" because of its arbitrary nature.
- A *Fortune* article stated that "of 58 large companies that have announced Six Sigma programs, 91 percent have trailed (performed below) the S&P 500 index since"
- Freeman Dyson (a famous theoretical physicist) once sat on a committee reviewing the Department of Energy Joint Genomics Institute (DOE JGI)
- Motorola sent their six-sigma preacher Freeman Dyson asked him:
	- • D: Can you explain me what is six–sigma?
		- • P: Mumbling something about it being the gold standard of reliability
	- • D: Can you at least define one-sigma?
		- •P: Silence
- • Six-sigma was never implemented at JGI



Born: December 15, 1923, Crowthorne, UK Died: February 28, 2020 Princeton, NJ USA

# Dyson's legacy

- **Seminal contributions to quantum mechanics**
- The Origin of Life: Cells  $\rightarrow$  Enzymes  $\rightarrow$  DNA/RNA later First proposed by Alexander Oparin in 1922
- Dyson sphere: **Completely** captures light from a star
- Dyson tree: genetically engineered tree growing inside a comet

