ASEN 5070
Statistical Orbit determination I
Fall 2012

Colorado Center for
Astrodynamics Research

Professor George H. Born
Professor Jeffrey S. Parker

Lecture 5: Stat OD

Announcements

- Homework 2 due Thursday

- Homework 3 out today
  - Basic dynamical systems relationships
  - Studies of the state transition matrix
  - Linear algebra

- I'm unavailable this Wednesday. Use those TAs ☺ and email is great of course.

Quiz Results

CCAR
Colorado Center for
Astrodynamics Research

Question 1 (1 point)

Are the following two time-series correlated? (more than just by chance. i.e., do they have an R-value greater than 0.5)

TIME - VALUE 1 - VALUE 2

1	6.3	-19.3
2	16.1	-17.8
3	18.9	-10.9
4	27.1	-9.4
5	30.9	-1.4
6	36.8	-1.8
7	45.1	4.8
8	52.3	10.9
9	57.1	11.6
10	62.4	18.6

- True
- False

Quiz Results

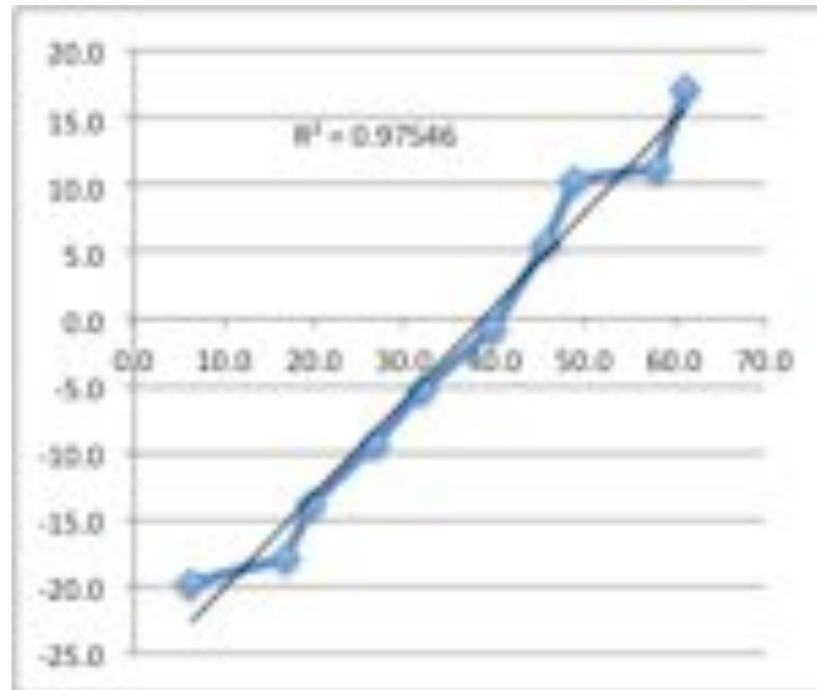
CCAR
Colorado Center for
Astrodynamics Research

Question 1 (1 point)

Are the following two time-series correlated? (more than just by chance. i.e., do they have an R-value greater than 0.5)

TIME	VALUE 1	VALUE 2
1	6.3	-19.3
2	16.1	-17.8
3	18.9	-10.9
4	27.1	-9.4
5	30.9	-1.4
6	36.8	-1.8
7	45.1	4.8
8	52.3	10.9
9	57.1	11.8
10	62.4	18.8

- True
- False



Quiz Results

CCAR
Colorado Center for
Astrodynamics Research

Question 2 (1 point)

If x and y are independent random variables drawn from Gaussian distributions with mean 0 and standard deviation 1, and f = x + y, what kind of distribution best describes f ?

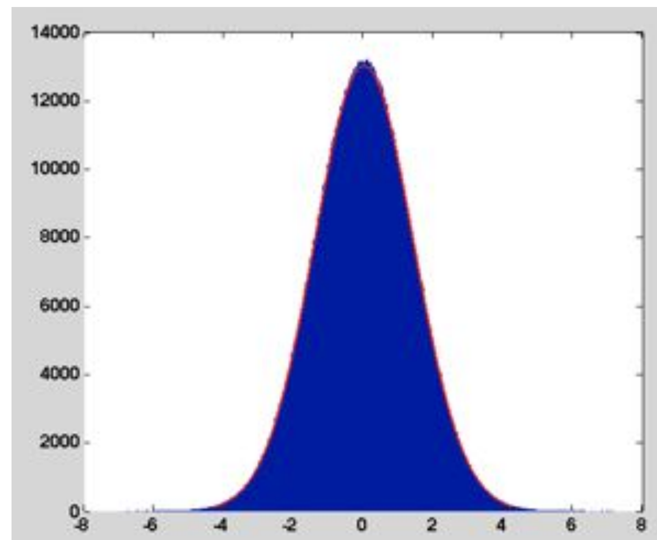
- A uniform distribution between 0 and 1
- A uniform distribution between 0 and 2
- A Gaussian distribution with mean 0.0
- A Gaussian distribution with mean 1.0

Quiz Results

Question 2 (1 point)

If x and y are independent random variables drawn from Gaussian distributions with mean 0 and standard deviation 1, and f = x + y, what kind of distribution best describes f ?

- A uniform distribution between 0 and 1
- A uniform distribution between 0 and 2
- A Gaussian distribution with mean 0.0
- A Gaussian distribution with mean 1.0



~N(0.0, 1.41)

Quiz Results

Question 3 (1 point)

If x and y are independent random variables drawn from uniform distributions between 0 and 1, and $f = x + y$, which of the following distributions does the best at describing f (albeit not perfectly).

- ○ A uniform distribution between 0 and 1
- ○ A uniform distribution between 0 and 2
- ○ A Gaussian distribution with mean 0.5
- ○ A Gaussian distribution with mean 1.0

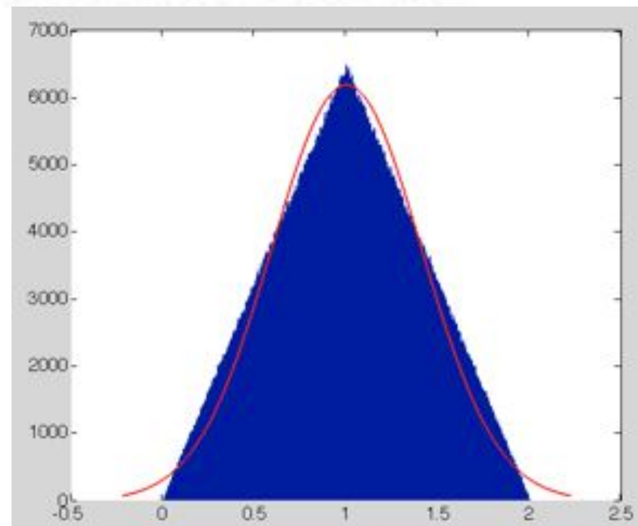
Quiz Results

CCAR
Colorado Center for Astrodynamics Research

Question 3 (1 point)

If x and y are independent random variables drawn from uniform distributions between 0 and 1, and f = x + y, which of the following distributions does the best at describing f (albeit not perfectly).

- A uniform distribution between 0 and 1
- A uniform distribution between 0 and 2
- A Gaussian distribution with mean 0.5
- A Gaussian distribution with mean 1.0



~N(1.0, 0.41)

Quiz Results

Question 4 (1 point)

If x is the number of people who fall asleep during an average StatOD lecture and "x" is drawn from a uniform distribution between 0 and 2, then what is the expected value for the total number of instances of people falling asleep after 30 independent lectures? (I know, only two at most? Unlikely.)

- 0
- 15
- 30
- 60

Quiz Results

CCAR
Colorado Center for Astrodynamics Research

Question 4 (1 point)

If x is the number of people who fall asleep during an average StatOD lecture and "x" is drawn from a uniform distribution between 0 and 2, then what is the expected value for the total number of instances of people falling asleep after 30 independent lectures? (I know, only two at most? Unlikely.)

- 0
- 15
- 30
- 60

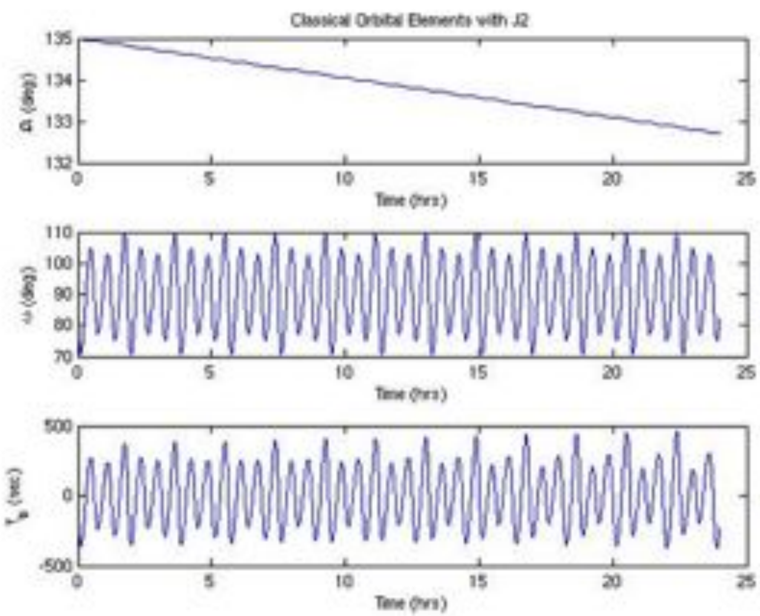
Homework 2

- Some popular questions and answers
- Energy with Drag

Homework 2

- Some popular questions and answers

- Computation of Time of Perigee



Homework 2: Tp Calculation

T_p is determined from the following equations:

$$M = E - e\sin E \qquad n = \sqrt{\frac{\mu}{a^3}} \qquad t - T_p = M/n$$

However, as time t increases, T_p is not constrained to an orbital period and thus increases as a step function. To resolve this, MOD T_p with the orbital period.

$$T_p = \mathrm{MOD}(T_p, P)$$

A situation may arise in which the calculation for the mean Anomaly, M, and true anomaly, v, do not agree resulting in the mean anomaly to be past perigee while the true anomaly is behind perigee (this is an artifact of numerical integration).

Homework 2: Tp Calculation

To correct this, we will introduce the angle of periapse θ_p:

$$\theta_p = nT_p$$

From this, one will notice that the artifacts occur when

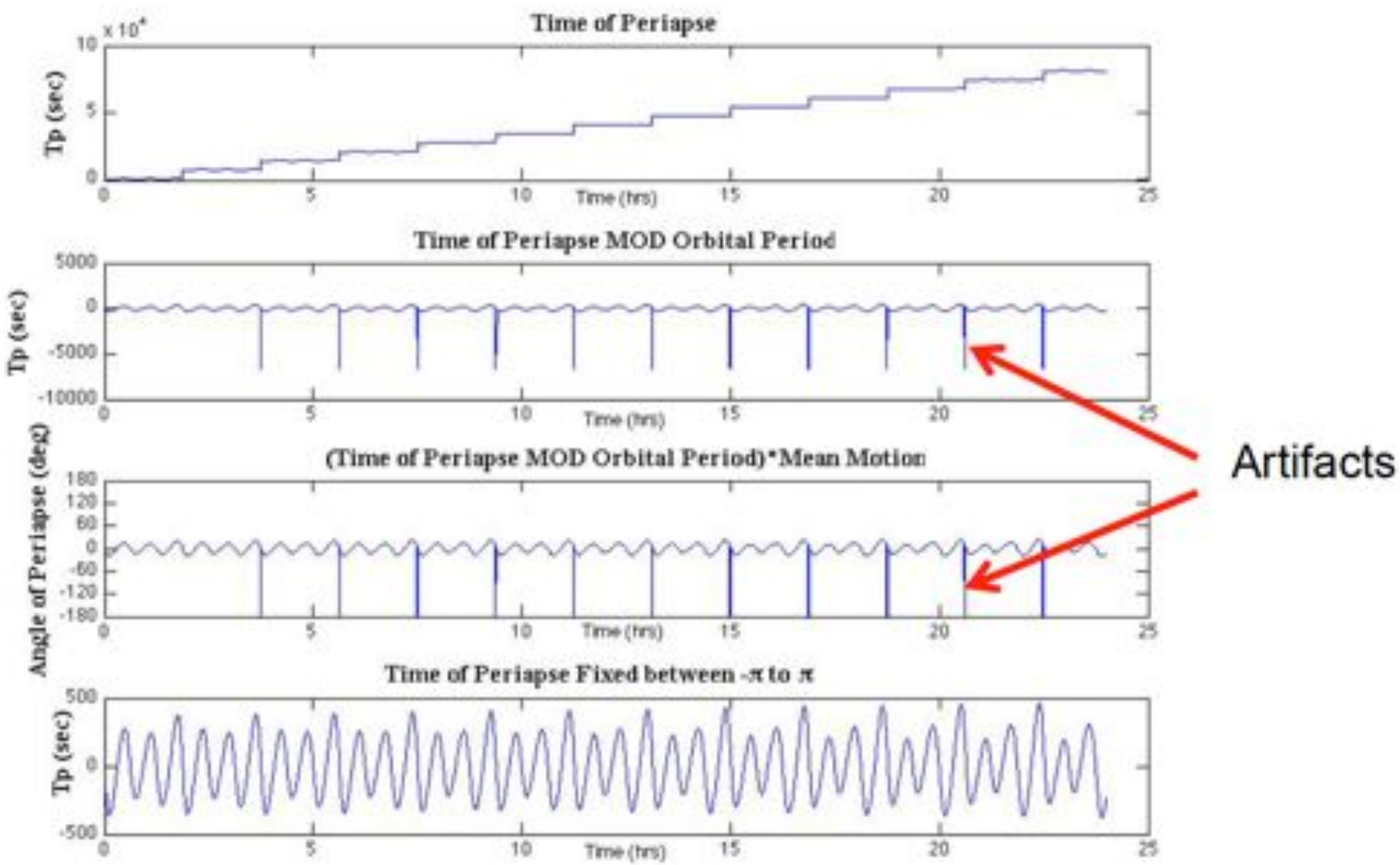
$$-\pi < \theta_p \leq \pi$$

Thus, constraining θ_p to be between $-\pi$ to π will remove the artifacts. The angle of periapse θ_p can then be converted back to time of periapse T_p by

$$T_p = \frac{\theta_p}{n}$$

Homework 2: Tp Calculation

CCAR
Colorado Center for Astrodynamics Research



Homework 3

- Chapter 4, Problems 1–6
  - Solving ODEs
  - Linear Algebra
  - Studying the state transition matrix

Today's Lecture

- Review of Differential Equations
  - Laplace Transforms
- Review of Statistics

Review of Diff EQ

- Stat OD dynamics:

$$\dot{\Phi} = A\Phi$$

- Solve for Φ given *A* and $\Phi(t_0, t_0) = I$

Review of Diff EQ

CCAR
Colorado Center for Astrodynamics Research

- Stat OD dynamics:

$$\dot{\Phi} = A\Phi$$

- Solve for Φ given *A* and $\Phi(t_0, t_0) = I$

$$A = \begin{bmatrix} a & 0 \\ b & g \end{bmatrix}$$

$$\begin{bmatrix} \dot{\phi}_{11} & \dot{\phi}_{12} \\ \dot{\phi}_{21} & \dot{\phi}_{22} \end{bmatrix} = \begin{bmatrix} a & 0 \\ b & g \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$$

Review of Diff EQ

$$\begin{bmatrix} \dot{\phi}_{11} & \dot{\phi}_{12} \\ \dot{\phi}_{21} & \dot{\phi}_{22} \end{bmatrix} = \begin{bmatrix} a & 0 \\ b & g \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$$

- Solve for ϕ_{11}
- w/ $\phi_{11}(t_0) = 1$

$$\dot{\phi}_{11} = a\phi_{11}$$

Review of Diff EQ

CCAR
Colorado Center for Astrodynamics Research

$$\begin{bmatrix} \dot{\phi}_{11} & \dot{\phi}_{12} \\ \dot{\phi}_{21} & \dot{\phi}_{22} \end{bmatrix} = \begin{bmatrix} a & 0 \\ b & g \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$$

- Solve for ϕ_{11}
- w/ $\phi_{11}(t_0) = 1$

$$\dot{\phi}_{11} = a\phi_{11}$$

$$\phi_{11}(t) = ce^{a(t-t_0)}$$
$$\phi_{11}(t_0) = 1$$
$$\phi_{11}(t) = e^{a(t-t_0)}$$

Example

- Solve the ODE

$$\dot{y} = ay$$

- We can solve this using "traditional" calculus:

$$\frac{dy}{dt} = ay$$

$$\frac{dy}{y} = adt$$

$$\ln y = a(t - t_0) + c$$

$$y = e^{a(t-t_0)+c}$$

$$y = c_1 e^{a(t-t_0)}$$

Check your answer by plugging it back in

Laplace Transforms

- Laplace Transforms are useful for analysis of linear time-invariant systems:
  - electrical circuits,
  - harmonic oscillators,
  - optical devices,
  - mechanical systems,
  - even orbit problems.

- Transformation from the time domain into the frequency domain.

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t) dt$$

- Inverse Laplace Transform converts the system back.

$$f(t) = \mathcal{L}^{-1}\{F(s)\} = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{\gamma-iT}^{\gamma+iT} e^{st} F(s) ds$$

Laplace Transform Tables

CCAR
Colorado Center for Astrodynamics Research

Function	Time domain $f(t) = \mathcal{L}^{-1}\{F(s)\}$	Laplace s-domain $F(s) = \mathcal{L}\{f(t)\}$	Region of convergence
unit impulse	$\delta(t)$	1	all s
delayed impulse	$\delta(t-\tau)$	$e^{-\tau s}$	
unit step	$u(t)$	$\frac{1}{s}$	$\mathrm{Re}\{s\} > 0$
delayed unit step	$u(t-\tau)$	$\frac{e^{-\tau s}}{s}$	$\mathrm{Re}\{s\} > 0$
ramp	$t \cdot u(t)$	$\frac{1}{s^2}$	$\mathrm{Re}\{s\} > 0$
delayed *n*th power with frequency shift	$\frac{(t-\tau)^n}{n!} e^{-\alpha(t-\tau)} \cdot u(t-\tau)$	$\frac{e^{-\tau s}}{(s+\alpha)^{n+1}}$	$\mathrm{Re}\{s\} > -\alpha$
*n*th power (for integer *n*)	$\frac{t^n}{n!} \cdot u(t)$	$\frac{1}{s^{n+1}}$	$\mathrm{Re}\{s\} > 0$ $(n > -1)$
*q*th power (for complex *q*)	$t^q \cdot u(t)$	$\frac{\Gamma(q+1)}{s^{q+1}}$	$\mathrm{Re}(s) > 0$ $\mathrm{Re}(q) > -1$
*n*th root	$\sqrt[n]{t} \cdot u(t)$	$\frac{\Gamma(\frac{1}{n}+1)}{s^{\frac{1}{n}+1}}$	$\mathrm{Re}\{s\} > 0$
*n*th power with frequency shift	$\frac{t^n}{n!} e^{-\alpha t} \cdot u(t)$	$\frac{1}{(s+\alpha)^{n+1}}$	$\mathrm{Re}\{s\} > -\alpha$
exponential decay	$e^{-\alpha t} \cdot u(t)$	$\frac{1}{s+\alpha}$	$\mathrm{Re}\{s\} > -\alpha$

exponential decay	$e^{-\alpha t} \cdot u(t)$	$\frac{1}{s+\alpha}$	$\mathrm{Re}\{s\} > -\alpha$
two-sided exponential decay	$e^{-\alpha\|t\|}$	$\frac{2\alpha}{\alpha^2 - s^2}$	$-\alpha < \mathrm{Re}\{s\} < \alpha$
exponential approach	$(1-e^{-\alpha t}) \cdot u(t)$	$\frac{\alpha}{s(s+\alpha)}$	$\mathrm{Re}\{s\} > 0$
sine	$\sin(\omega t) \cdot u(t)$	$\frac{\omega}{s^2+\omega^2}$	$\mathrm{Re}\{s\} > 0$
cosine	$\cos(\omega t) \cdot u(t)$	$\frac{s}{s^2+\omega^2}$	$\mathrm{Re}\{s\} > 0$
hyperbolic sine	$\sinh(\alpha t) \cdot u(t)$	$\frac{\alpha}{s^2-\alpha^2}$	$\mathrm{Re}\{s\} > \|\alpha\|$
hyperbolic cosine	$\cosh(\alpha t) \cdot u(t)$	$\frac{s}{s^2-\alpha^2}$	$\mathrm{Re}\{s\} > \|\alpha\|$
Exponentially decaying sine wave	$e^{-\alpha t}\sin(\omega t) \cdot u(t)$	$\frac{\omega}{(s+\alpha)^2+\omega^2}$	$\mathrm{Re}\{s\} > -\alpha$
Exponentially decaying cosine wave	$e^{-\alpha t}\cos(\omega t) \cdot u(t)$	$\frac{s+\alpha}{(s+\alpha)^2+\omega^2}$	$\mathrm{Re}\{s\} > -\alpha$
natural logarithm	$\ln(t) \cdot u(t)$	$-\frac{1}{s}[\ln(s)+\gamma]$	$\mathrm{Re}\{s\} > 0$
Bessel function of the first kind, of order *n*	$J_n(\omega t) \cdot u(t)$	$\frac{\left(\sqrt{s^2+\omega^2}-s\right)^n}{\omega^n\sqrt{s^2+\omega^2}}$	$\mathrm{Re}\{s\} > 0$ $(n > -1)$
Error function	$\mathrm{erf}(t) \cdot u(t)$	$\frac{e^{s^2/4}(1-\mathrm{erf}(s/2))}{s}$	$\mathrm{Re}\{s\} > 0$

$$\mathcal{L}\{f'(t)\} = s\mathcal{L}\{f(t)\} - f(t_0)$$

Example

- Solve the ODE

$$\dot{y} = ay$$

- We can solve this using "traditional" calculus:

$$\begin{aligned}
\frac{dy}{dt} &= ay \\
\frac{dy}{y} &= a dt \\
\ln y &= a(t - t_0) + c \\
y &= e^{a(t-t_0)+c} \\
y &= c_1 e^{a(t-t_0)}
\end{aligned}$$

Example

- Solve the ODE

$$\dot{y} = ay$$

- Or, we can solve this using Laplace Transforms:

$$\frac{dy}{dt} = ay$$

$$\mathcal{L}\left\{\frac{dy}{dt}\right\} = \mathcal{L}\left\{ay\right\}$$

$$s\mathcal{L}\left\{y\right\} - y(t_0) = a\mathcal{L}\left\{y\right\}$$

$$\mathcal{L}\left\{y\right\} = \frac{y(t_0)}{s-a}$$

$$\begin{aligned} y &= y(t_0)\mathcal{L}^{-1}\left\{\frac{1}{s-a}\right\} \\ &= y(t_0)e^{a(t-t_0)} \end{aligned}$$

Applied to Stat OD

- Solve the ODE $\dot{\Phi} = A\Phi$

Applied to Stat OD

- Solve the ODE $\dot{\Phi} = A\Phi$

$$\mathcal{L}\left\{\dot{\Phi}\right\} = \mathcal{L}\left\{A\Phi\right\}$$

Applied to Stat OD

- Solve the ODE $\dot{\Phi} = A\Phi$

$$\mathcal{L}\{\dot{\Phi}\} = \mathcal{L}\left\{A\Phi\right\}$$

$$s\mathcal{L}\left\{\Phi\right\} - \Phi_0 = A\mathcal{L}\left\{\Phi\right\}$$

Applied to Stat OD

- Solve the ODE $\dot{\Phi} = A\Phi$

$$\mathcal{L}\{\dot{\Phi}\} = \mathcal{L}\{A\Phi\}$$

$$s\mathcal{L}\{\Phi\} - \Phi_0 = A\mathcal{L}\{\Phi\}$$

$$s\mathcal{L}\{\Phi\} - A\mathcal{L}\{\Phi\} = \Phi_0$$

Applied to Stat OD

- Solve the ODE $\dot{\Phi} = A\Phi$

$$\mathcal{L}\{\dot{\Phi}\} = \mathcal{L}\{A\Phi\}$$

$$s\mathcal{L}\{\Phi\} - \Phi_0 = A\mathcal{L}\{\Phi\}$$

$$s\mathcal{L}\{\Phi\} - A\mathcal{L}\{\Phi\} = \Phi_0$$

$$\mathcal{L}\{\Phi\}(sI - A) = I$$

Applied to Stat OD

- Solve the ODE $\dot{\Phi} = A\Phi$

$$\mathcal{L}\{\dot{\Phi}\} = \mathcal{L}\left\{A\Phi\right\}$$

$$s\mathcal{L}\left\{\Phi\right\} - \Phi_0 = A\mathcal{L}\left\{\Phi\right\}$$

$$s\mathcal{L}\left\{\Phi\right\} - A\mathcal{L}\left\{\Phi\right\} = \Phi_0$$

$$\mathcal{L}\left\{\Phi\right\}\left(sI - A\right) = I$$

$$\mathcal{L}\left\{\Phi\right\} = \left(sI - A\right)^{-1}$$

Applied to Stat OD

- Solve the ODE $\dot{\Phi} = A\Phi$

$$\mathcal{L}\{\dot{\Phi}\} = \mathcal{L}\left\{A\Phi\right\}$$

$$s\mathcal{L}\left\{\Phi\right\} - \Phi_0 = A\mathcal{L}\left\{\Phi\right\}$$

$$s\mathcal{L}\left\{\Phi\right\} - A\mathcal{L}\left\{\Phi\right\} = \Phi_0$$

$$\mathcal{L}\left\{\Phi\right\}(sI - A) = I$$

$$\mathcal{L}\left\{\Phi\right\} = (sI - A)^{-1}$$

$$\Phi = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}$$

Questions?

- Questions on Diff EQ?
- Quick Break
- Review of Statistics to follow

Review of Statistics

- X is a random variable with a prescribed domain.
- x is a realization of that variable.

- Example:
  - $0 < X < 1$

  - $x_1 = 0.232$
  - $x_2 = 0.854$
  - $x_3 = 0.055$
  - etc

Axioms of Probability

- $p(A)$ is the numerical probability of an event A occurring, e.g., A is the event that heads appears on the flip of coin.

- $p(AB)$ is the numerical probability that both events A and B occur.

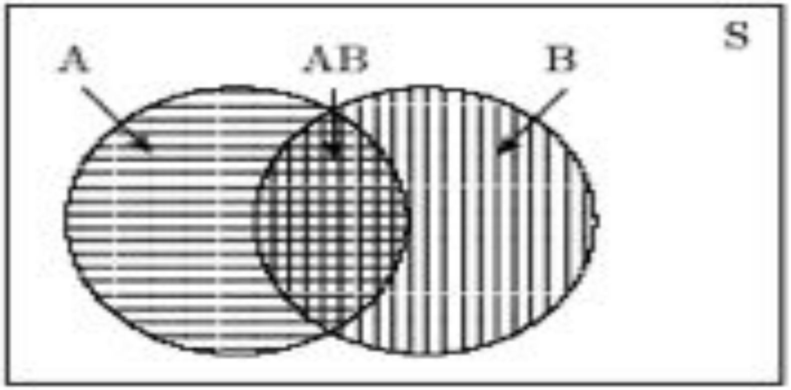
- The Axioms of Probability

    1. $p(A) \geq 0$

    2. p(S)=1, S is the certain event

    3. $p(A + B) = p(A) + p(B) - p(AB)$

Venn Diagram



S :	all points in the rectangle
A :	points in horizontal hatch region
B :	points in vertical hatch region
AB :	points in cross hatch region
$A+B$:	all points in hatched area

$$p(A+B) = p(A) + p(B) - p(AB). \qquad \text{(A.2.1)}$$

Axioms of Probability

- Conditional Probability

$$p\left(A/B\right) = \frac{p\left(AB\right)}{p\left(B\right)}$$

- Two events are independent if

$$p\left(A/B\right) = p\left(A\right) \text{ and } P\left(B/A\right) = P\left(B\right)$$

i.e., $$p\left(AB\right) = p\left(A\right)p\left(B\right)$$

Probability Density & Distribution Functions

Probability over the interval $x, x + dx$ is defined in

$$p(x \leq X \leq x + dx) = f(x)dx\,. \tag{A.4.1}$$

- For the continuous random variable, axioms 1 and 2 become

1. $f(x) \geq 0$ (A.4.2)
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1\,.$ (A.4.3)

Probability Density & Distribution Functions

Probability over the interval x, $x + dx$ is defined in

$$p(x \leq X \leq x + dx) = f(x)dx\,. \qquad \text{(A.4.1)}$$

- For the continuous random variable, axioms 1 and 2 become

1. $f(x) \geq 0$ (A.4.2)
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1\,.$ (A.4.3)

- The third axiom becomes

3. $p(a \leq X \leq c) = \int_a^c f(x)\,dx$ (A.4.4)

- Which for a < b < c

$$\begin{aligned} p(a \leq X \leq c) &= \int_a^b f(x)\,dx + \int_b^c f(x)\,dx \\ &= p(a \leq X \leq b) + p(b \leq X \leq c)\,. \end{aligned} \qquad \text{(A.4.5)}$$

Probability Density & Distribution Functions

2. $$\int_{-\infty}^{\infty} f(x)\,dx = 1\,. \tag{A.4.3}$$

Using axiom 2 as a guide, solve the following for *k*:

$$f(x) = \left\{ \begin{array}{cc} k(x+1) & 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{array} \right\}$$

Probability Density & Distribution Functions

Colorado Center for Astrodynamics Research

2. $$\int_{-\infty}^{\infty} f(x)\, dx = 1\,. \qquad \text{(A.4.3)}$$

Using axiom 2 as a guide, solve the following for *k*:

$$f(x) = \left\{ \begin{array}{cc} k(x+1) & 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{array} \right\}$$

$$k \int_0^1 (x+1)dx = 1$$

$$k \left[\frac{x^2}{2} + x \right]_0^1 = 1 \qquad k\frac{3}{2} = 1 \rightarrow k = \frac{2}{3}$$

Probability Density & Distribution Functions

Out of interest in the event $X \leq x$, we introduce $F(x)$, the *distribution function* of the continuous random variable X, and define it by

$$F(x) = p(X \leq x) = \int_{-\infty}^{x} f(t)\, dt. \qquad \text{(A.4.6)}$$

It follows that

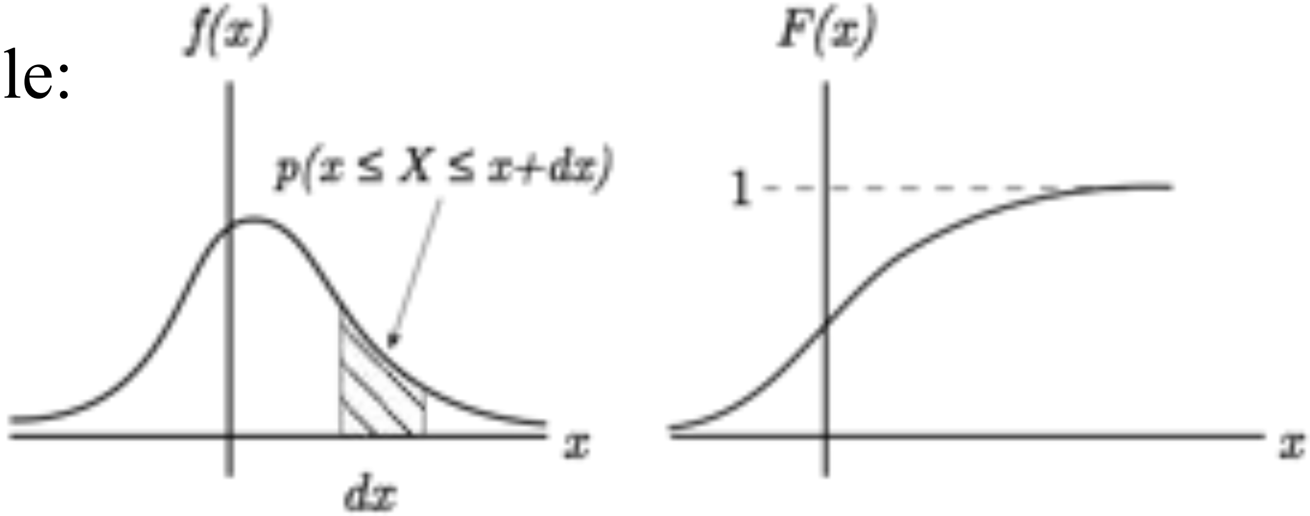
$$F(-\infty) = 0 \text{ and } F(\infty) = 1\,.$$

From elementary calculus, at points of continuity of F

$$\frac{d\,F(x)}{dx} = f(x) \qquad \text{(A.4.7)}$$

Probability Density & Distribution Functions

Example:



Density function of a continuous random variable

Distribution function of a continuous random variable

- From the definition of the density and distribution functions we have:

$$p\,(a \leq X \leq b) = \int_a^b f\,(x)\,dx = F\,(b) - F\,(a)\,. \tag{A.4.8}$$

- From axioms 1 and 2, we find:

$$0 \leq F(x) \leq 1 \tag{A.4.9}$$

Expected Values

The *expected value* or the *mean* of X is written $E\left(X\right)$, and is defined by

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx\,. \tag{A.5.1}$$

- The kth moment of X about the origin is

$$E\left[X^k\right] = \lambda_k = \int\limits_{-\infty}^{\infty} x^k f(x)\, dx\,. \tag{A.6.1}$$

Note that:

$$\lambda_0 = \int_{\infty}^{\infty} f(x) dx = 1$$

$$\lambda_1 = \int_{\infty}^{\infty} x f(x) dx = \text{Mean}$$

$$\lambda_2 = \int_{\infty}^{\infty} x^2 f(x) dx$$

Expected Values

We also may speak of the kth moment of X about the mean λ_1. In this case, we define

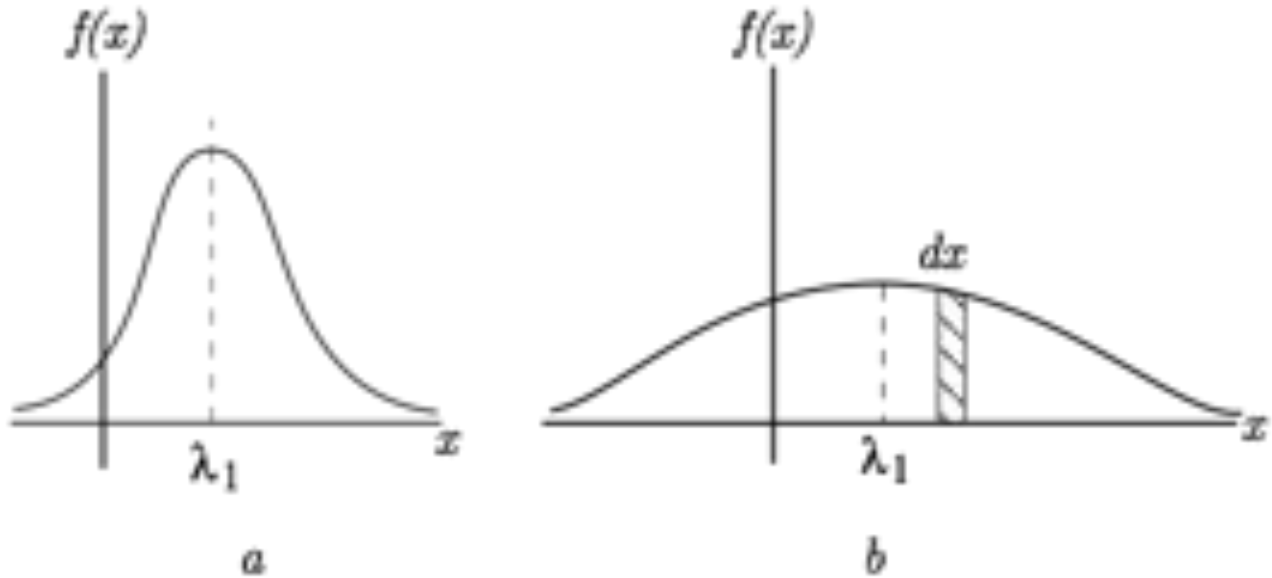
$$\mu_k \equiv E\,(X - \lambda_1)^k = \int_{-\infty}^{\infty} (x - \lambda_1)^k f(x)dx\,. \qquad \text{(A.6.2)}$$

- Note that μ_2 also denoted as $\sigma^2(X)$ is called the variance of x

$$\mu_2 = E\,(X - \lambda_1)^2 = \int_{-\infty}^{\infty} (x - \lambda_1)^2 f(x)dx\,. \qquad \text{(A.6.3)}$$

Expected Values

- σ^2 is one measure of the dispersion of the distribution about its mean value.



Both density functions *a* and *b* have their mean value indicated by λ_1. Note that the variance of *b* will be much larger than that of *a*.

Expected Values

- Also

$$\begin{aligned}\mu_2 &= \int_{-\infty}^{\infty} (x-\lambda_1)^2 f(x)\, dx \\ &= \int_{-\infty}^{\infty} (x^2 - 2x\lambda_1 + \lambda_1^2) f(x) dx \\ &= \lambda_2 - 2\lambda_1^2 + \lambda_1^2 \\ &= \lambda_2 - \lambda_1^2 . \end{aligned} \tag{A.6.6}$$

The higher order moments are of theoretical importance in any distribution, but they do not have a simple geometric or physical interpretation as do λ_1 and μ_2.

Expected Values

- A few useful results follow readily from the definition of the expected value and the fact that it is a linear operator

$$E(a + bX) = a + bE(X) \tag{A.6.4}$$

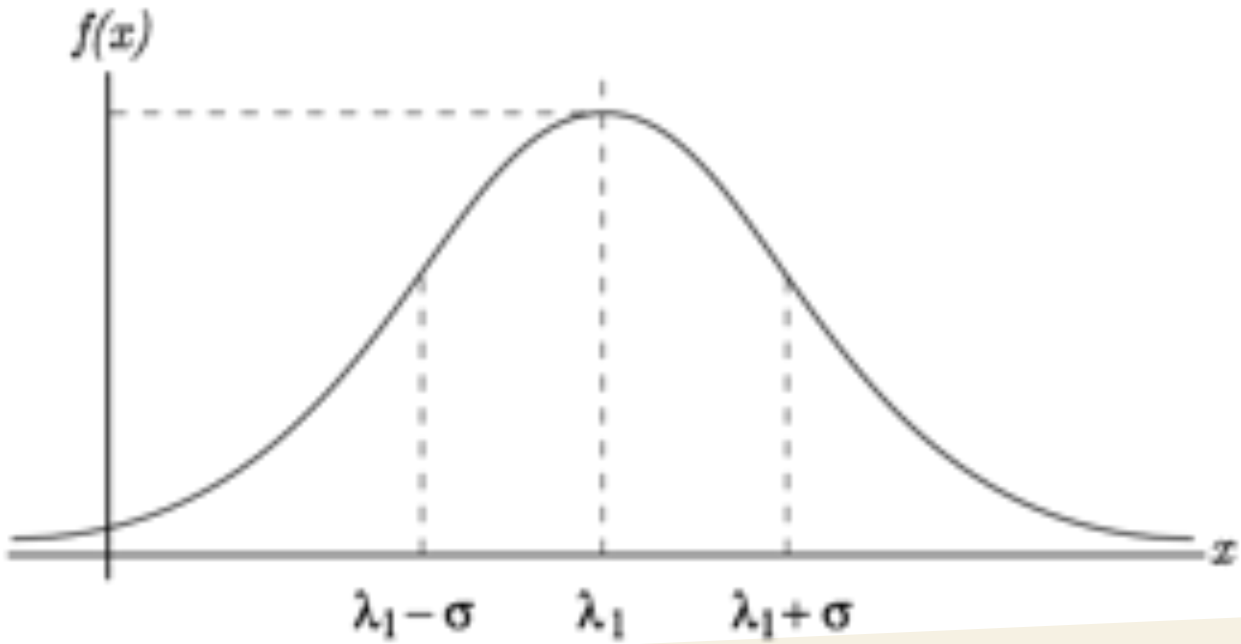
where a and b are constants.

The Gaussian or Normal Density Function

One of the most important density functions is the Gaussian.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\lambda_1}{\sigma}\right)^2\right] -\infty < x < \infty \,. \tag{A.8.7}$$

- The Gaussian density function is depicted graphically as



The Gaussian or Normal Density Function

- Other properties of the univariate normal distribution function are

$$p(\lambda_1 - \sigma \leq X \leq \lambda_1 + \sigma) = \int_{\lambda_1 - \sigma}^{\lambda_1 + \sigma} f(x)dx = .68268$$

$$p\left[(\lambda_1 - 2\sigma) \leq X \leq (\lambda_1 + 2\sigma)\right] = .95449$$

$$p\left[(\lambda_1 - 3\sigma) \leq X \leq (\lambda_1 + 3\sigma)\right] = .99730.$$

Moment Generating Functions

Consider the particular case of

$$E\left[g(X)\right] = \int_{-\infty}^{\infty} g(x)f(x)dx \qquad \text{(A.7.1)}$$

for which

$$g(X) = e^{\theta x}$$

where θ is a dummy variable. Since

$$e^{\theta x} = 1 + \theta x + \frac{(\theta x)^2}{2!} + \cdots \frac{(\theta x)^n}{n!} + \cdots \qquad \text{(A.7.2)}$$

substituting Eq. (A.7.2) into Eq. (A.7.1) results in

$$E\left(e^{\theta x}\right) = \lambda_0 + \theta\lambda_1 + \frac{\theta^2\lambda_2}{2!} + \ \ldots\frac{\theta^n\lambda_n}{n!} + \cdots \qquad \text{(A.7.3)}$$

Moment Generating Functions

Thus $E(e^{\theta x})$ may be said to generate the moments $\lambda_0, \lambda_1 \ldots \lambda_n$ of the random variable X. It is called the *moment generating function* of X and is written $M_X(\theta)$. Note that

$$\left.\frac{\partial^k M_X(\theta)}{\partial \theta^k}\right|_{\theta=0} = \lambda_k. \tag{A.7.4}$$

Accepting the fact that the moment generating function for the function $h(X)$ is given by

$$M_{h(X)}(\theta) = \int_{-\infty}^{\infty} e^{\theta h(X)} f(x)dx, \tag{A.7.5}$$

let $h(X) = X - \lambda_1$, then

$$M_{(X-\lambda_1)}(\theta) = e^{-\theta\lambda_1} M_X(\theta) \tag{A.7.6}$$

which relates moments about the origin to moments about the mean,

$$\mu_k = \left.\frac{\partial^k M_{(X-\lambda_1)}(\theta)}{\partial \theta^k}\right|_{\theta=0}. \tag{A.7.7}$$

Moment Generating Functions

From Eqs. (A.7.3) and (A.7.6)

$$M_{(X-\lambda_1)}(\theta) = e^{-\theta\lambda_1}\left(\lambda_0 + \theta\lambda_1 + \frac{\theta^2\lambda_2}{2!} + \cdots \frac{\theta^n\lambda_n}{n!}\cdots\right)$$

and for example,

$$\begin{aligned}
\mu_2 &= \left.\frac{\partial^2 M_{(X-\lambda_1)}(\theta)}{\partial\theta^2}\right|_{\theta=0} \\
&= \lambda_1^2 e^{-\theta\lambda_1}(\lambda_0 + \theta\lambda_1) - \lambda_1 e^{-\theta\lambda_1}\lambda_1 - \lambda_1 e^{-\theta\lambda_1}\lambda_1 + e^{-\theta\lambda_1}\lambda_2\Big|_{\theta=0} \\
&= \lambda_2 - \lambda_1^2, \quad \text{recall that} \quad \lambda_0 = 1\,.
\end{aligned}$$

Two Random Variables

In analogy with a single random variable, the density function for two random variables may be written as

$$p\,(x \leq X \leq x + dx, y \leq Y \leq y + dy) = f(x, y)dxdy\,. \qquad \text{(A.9.6)}$$

- In summary, for two random variables

$F\,(x,\ y)$	$\equiv$	joint distribution function of $X,\ Y$.
$f\,(x,\ y)$	$\equiv$	joint density function of $X,\ Y$.
$f\,(x,\ y)\,dx\,dy$	$\equiv$	joint probability element of $X,\ Y$.

Marginal Distributions

We often want to determine the probability behavior of one random variable, given the joint probability behavior of two. This is interpreted to mean

$$p\left(X \leq x, \text{no condition on } Y\right) = F\left(x, \infty\right). \tag{A.10.1}$$

where

$$\begin{aligned} F(x,\infty) &= \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(u,v)\, dv\, du = \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f(u,v)\, dv \right] du \\ &= \int_{-\infty}^{x} g\left(u\right) du\,. \end{aligned} \tag{A.10.2}$$

Marginal Distributions

- and

$$g\,(x) = \int_{-\infty}^{\infty} f(x, y)\,dy \tag{A.10.3}$$

is called the marginal density function of X.

- Similarly,

$$h\,(y) = \int_{-\infty}^{\infty} f\,(x,\,y)dx \tag{A.10.4}$$

is the marginal density function of Y

Independence of Random Variables

X and Y are independent if we can factor their joint density function into

$$f(x, y) = g(x)h(y) \tag{A.11.1}$$

where $g(x)$ and $h(y)$ are the marginal density functions of X and Y.

Conditional Probability

CCAR
Colorado Center for Astrodynamics Research

In analogy with the simple events A and B, where we define $p\left({}^{B}\!/_{A} \right)$ by

$$p(B/A) = \frac{p(AB)}{p(A)}. \tag{A.12.1}$$

We define a *conditional density function* for continuous random variables X and Y with density functions $f(x, y)$, $g(x)$, and $h(y)$ by

$$g(x/y) = \frac{f(x,y)}{h(y)}, \; h(y/x) = \frac{f(x,y)}{g(x)}. \tag{A.12.2}$$

- If X and Y are independent we have

$$g(x/y) = \frac{f(x,y)}{h(y)} = \frac{g(x)h(y)}{h(y)} = g(x)$$

(A.12.5)

$$h(y/x) = \frac{f(x,y)}{g(x)} = h(y).$$

Expected Values of Bivariate Functions

- The expected value $E\left[\phi\left(X, Y\right)\right]$ of an arbitrary function $\phi\left(X, Y\right)$ of two continuous random variables X and Y is given by

$$E\left[\phi\left(X, Y\right)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi\left(x, y\right) f\left(x, y\right) dx\, dy\,. \tag{A.13.1}$$

Expected Values of Bivariate Functions

As with one random variable, the expected value of certain functions is of great importance in identifying characteristics of joint probability distributions.

- Setting

$$\phi(X, Y) = X^l Y^m$$

yields

$$E[X^l Y^m] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^l y^m f(x, y) dx\, dy \equiv \lambda_{lm} \qquad \text{(A.13.2)}$$

written λ_{lm}, the lm^{th} moment of X, Y about the origin.

Expected Values of Bivariate Functions

- The lm^{th} moment about the mean is obtained by setting

$$\phi(X, Y) = [X - \lambda_{10}]^l [Y - \lambda_{01}]^m .$$

This results in

$$E\left\{[X - \lambda_{10}]^l\,[Y - \lambda_{01}]^m\right\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x - \lambda_{10}]^l [y - \lambda_{01}]^m f(x, y) dx\, dy \equiv \mu_{lm} \qquad \text{(A.13.3)}$$

Expected Values of Bivariate Functions

l	m	
0	0	$\lambda_{00} = 1$
1	0	$\lambda_{10} = E(X)$, the mean of X
0	1	$\lambda_{01} = E(Y)$, the mean of Y
0	0	$\mu_{00} = 1$
1	1	$\mu_{11} = E\{[X - E(X)](Y - E(Y))\}$, the covariance of X and Y
2	0	$\mu_{20} = \sigma^2(X)$, the variance of X
0	2	$\mu_{02} = \sigma^2(Y)$, the variance of Y.

Expected Values of Bivariate Functions

CCAR
Colorado Center for Astrodynamics Research

l	m	
0	0	$\lambda_{00} = 1$
1	0	$\lambda_{10} = E(X)$, the mean of X
0	1	$\lambda_{01} = E(Y)$, the mean of Y
0	0	$\mu_{00} = 1$
1	1	$\mu_{11} = E\{[X - E(X)](Y - E(Y))\}$, the covariance of X and Y
2	0	$\mu_{20} = \sigma^2(X)$, the variance of X
0	2	$\mu_{02} = \sigma^2(Y)$, the variance of Y.

Consider as an example the computation of

$$\begin{aligned}
\mu_{11} &= E(X - \lambda_{10})(Y - \lambda_{01}) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \lambda_{10})(y - \lambda_{01}) f(x, y) dx\, dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (xy - \lambda_{10} y - \lambda_{01} x + \lambda_{10}\lambda_{01}) f(x, y) dx\, dy \\
&= \lambda_{11} - 2\lambda_{10}\lambda_{01} + \lambda_{10}\lambda_{01} \\
&= \lambda_{11} - \lambda_{10}\lambda_{01}\,. \qquad \text{(A.13.4)}
\end{aligned}$$

The Variance-Covariance Matrix

The correlation coefficient is defined as

$$\rho_{XY} \equiv \frac{E\{[X - E(X)]\,[Y - E(Y)]\}}{\{E[X - E(X)]^2\}^{1/2}\{E\,[Y - E(Y)]^2\}^{1/2}} = \frac{\mu_{11}}{\sigma(X)\sigma(Y)}\,. \tag{A.14.2}$$

The variance-covariance matrix for an n-dimensional random vector, X, can be written as

$$P = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \rho_{2n}\sigma_2\sigma_n & \cdots & \sigma_n^2 \end{bmatrix} \tag{A.14.3}$$

where ρ_{ij} is a measure of the degree of linear correlation between X_i and X_j.

Properties of the Correlation Coefficient

CCAR
Colorado Center for Astrodynamics Research

- $-1 \leq \rho_{XY} \leq 1$

- If $\rho_{XY} = \pm 1$, there is a linear relationship between X and Y, i.e.,

$$X = a \pm bY$$

where a and b are constants

Properties of Covariance and Correlation

From the definition of μ_{11}

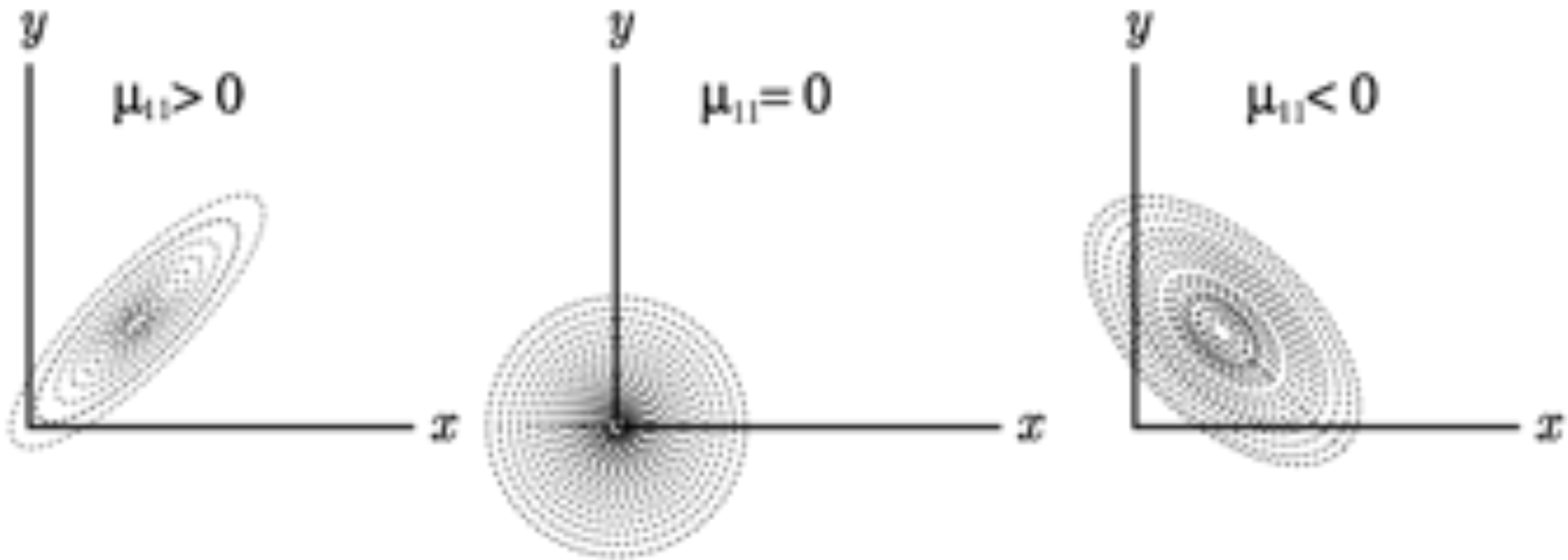
$$\mu_{11} = E[(X - E(X))(Y - E(Y))]$$

- If large values of X are found paired generally with large values of Y, and if small values of X are paired with small values of Y, μ_{11} and hence ρ_{XY} will be positive.

- If large values of X are paired with small values of Y, then μ_{11} and hence ρ_{XY} will be negative.

- If some large and small values of X and Y are paired then $\mu_{11} \simeq 0$.

Properties of Covariance and Correlation



- An example of positive correlation would be a sampling of human height and weight.

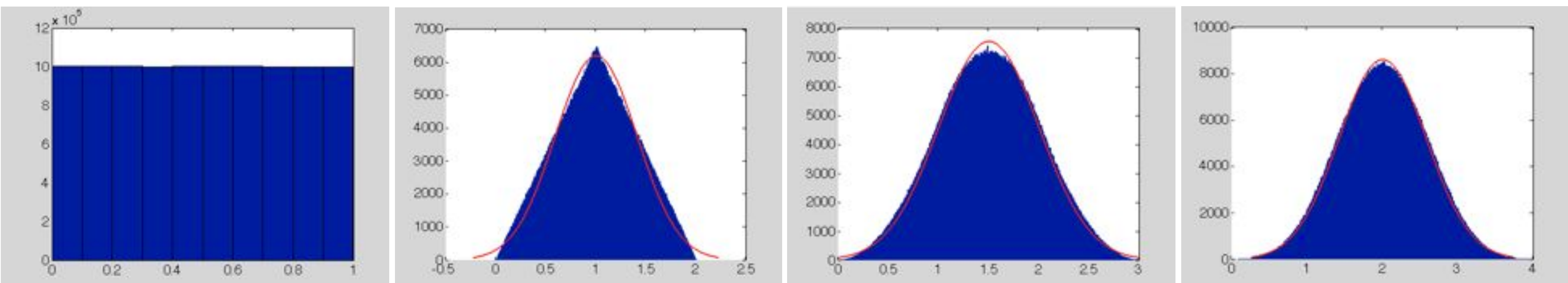
Central Limit Theorem

If a large number of sets of samples of size n are taken of a random variable, with any given density function, the mean of these samples will be normally distributed as $n \to \infty$.

Central Limit Theorem

- Addition of multiple variables taken from any single distribution → Gaussian
- Example: Uniform [0,1]

Sum of 1 var	Sum of 2 vars	Sum of 3 vars	Sum of 4 vars

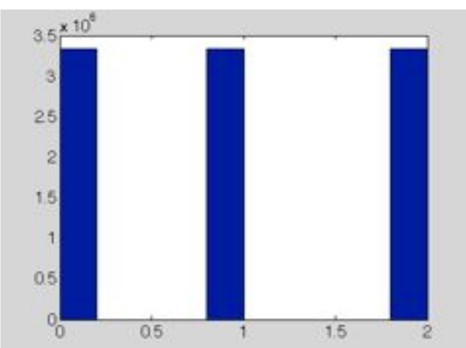


~N(0.5, 0.29)	~N(1.0, 0.41)	~N(1.5, 0.50)	~N(2.0, 0.58)

Central Limit Theorem

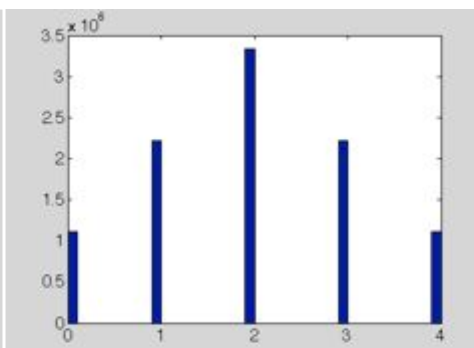
- Addition of multiple variables taken from any single distribution → Gaussian
- Example: Uniform {0,1,2} (Quiz Question #4)

Sum of 1 var



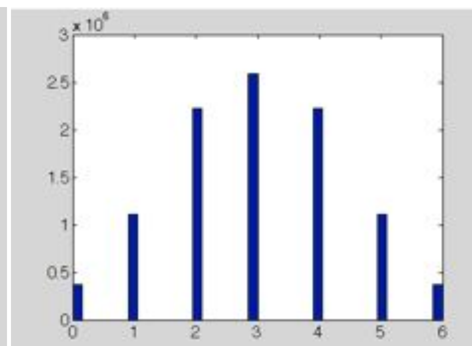
~N(1.0, 0.82)

Sum of 2 vars



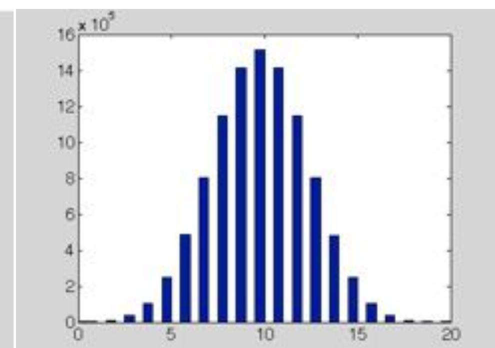
~N(2.0, 1.16)

Sum of 3 vars



~N(3.0, 1.42)

Sum of 10 vars

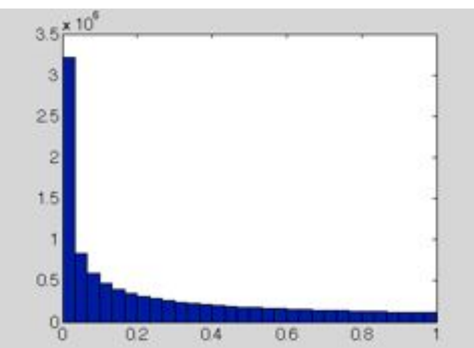
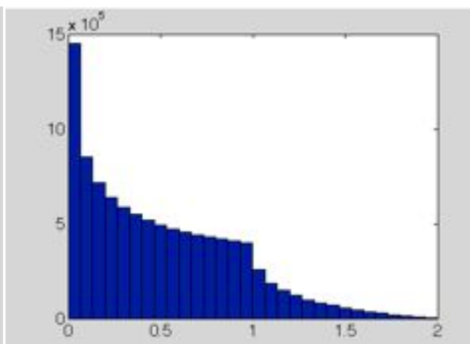
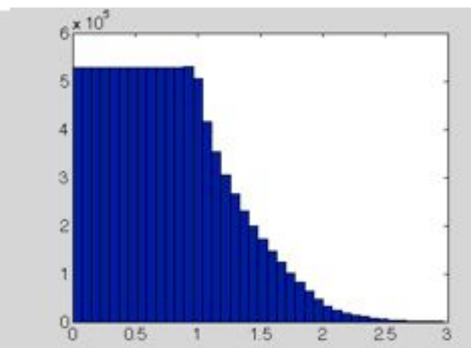
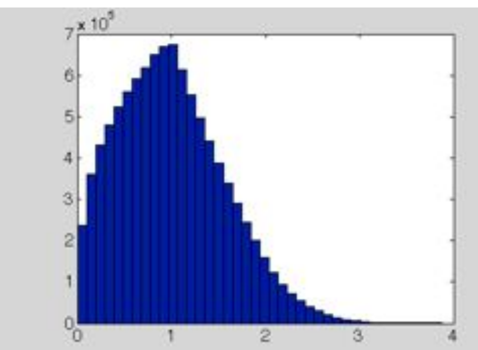


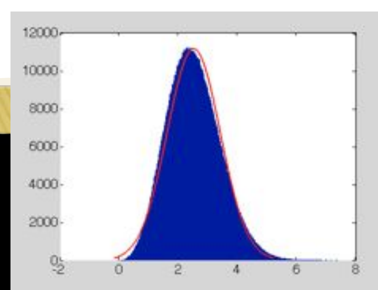
~N(10, 2.58)

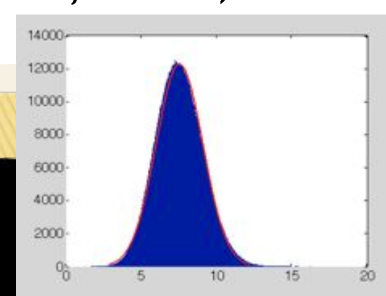
Central Limit Theorem

CCAR
Colorado Center for
Astrodynamics Research

- Addition of multiple variables taken from any single distribution → Gaussian
- Example: Skewed distribution

Sum of 1 var	Sum of 2 vars	Sum of 3 vars	Sum of 4 vars
			
~N(0.25, 0.28)	~N(0.5, 0.40)	~N(0.75, 0.49)	~N(1.0, 0.57)





Central limit theorem

52 Elementary Statistics

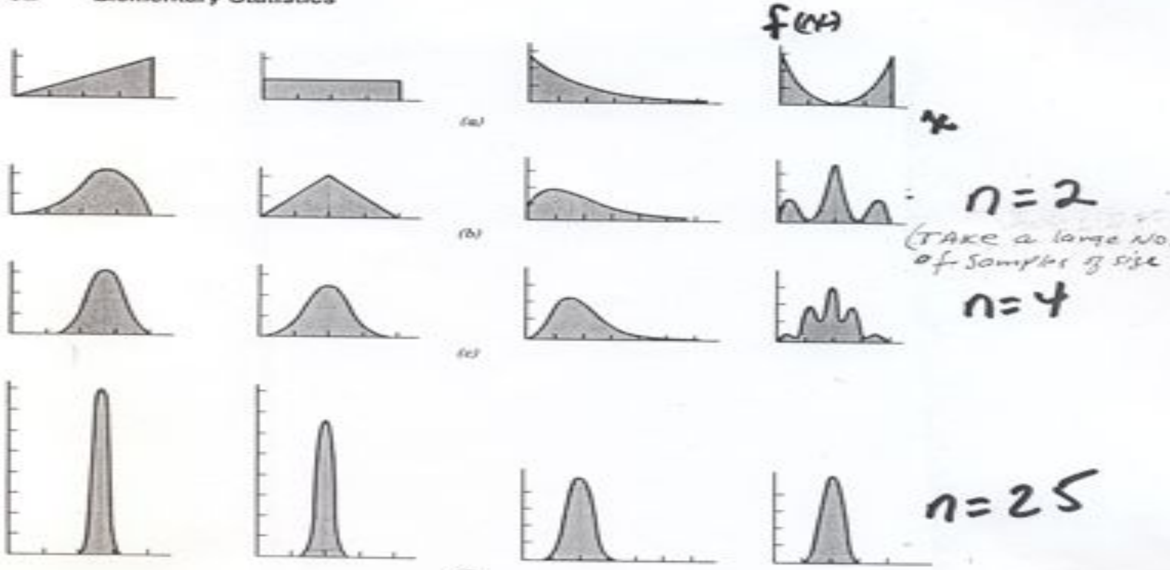


FIGURE 2.25 Distributions of $\overline{X}$ for large numbers of samples of size n, selected at random from nonnormal populations. Central limits theorem insures that distribution of $\overline{X}$ tends toward normal distribution as n increases. Parent populations from which samples are drawn are shown in (a). Distributions of $\overline{X}$ are shown for (b) samples of size $n = 2$, (c) samples of size $n = 4$, and (d) samples of size $n = 25$. After Lapin (1982).

From: Davis, J.C., Statistical Data analysis in Geology, John Wiley, 1986

Questions

- Questions on Statistics?
- I'll go through example problems at the beginning of Thursday's lecture
- Homework 2 due Thursday
- Homework 3 out today
- Next quiz active tomorrow at 1pm.