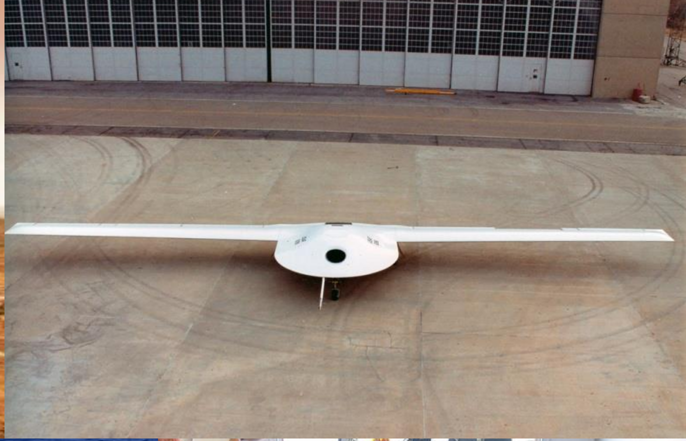
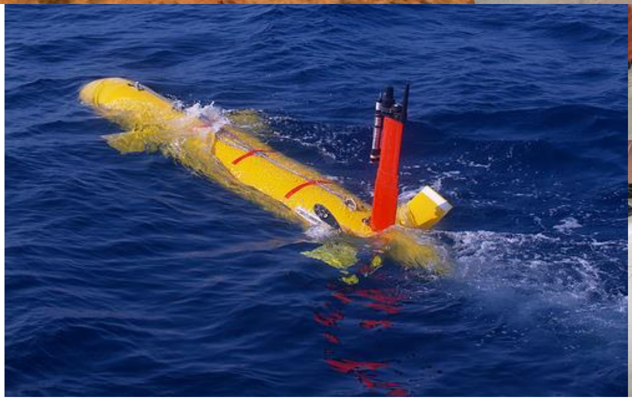
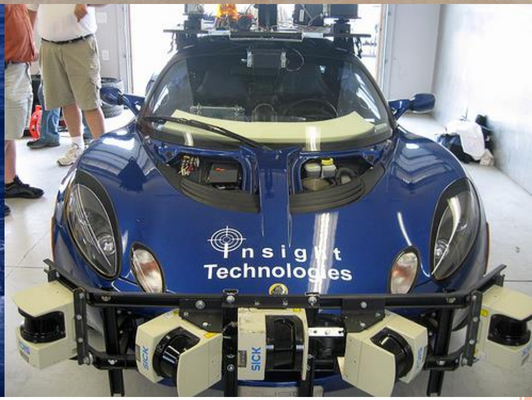




ME 597: Autonomous Mobile Robotics
Section 2 – Probability

Prof. Steven Waslander

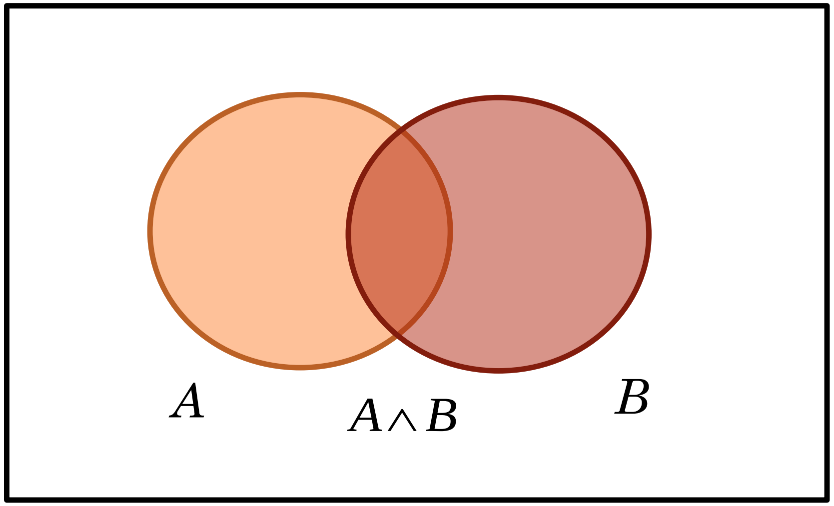
PROBABILITY

- *p(A)*: Probability that A is true

$$0 \leq p(A) \leq 1$$

$$p(True) = 1, p(False) = 0$$

$$p(A \vee B) = p(A) + p(B) - p(A \wedge B)$$



PROBABILITY

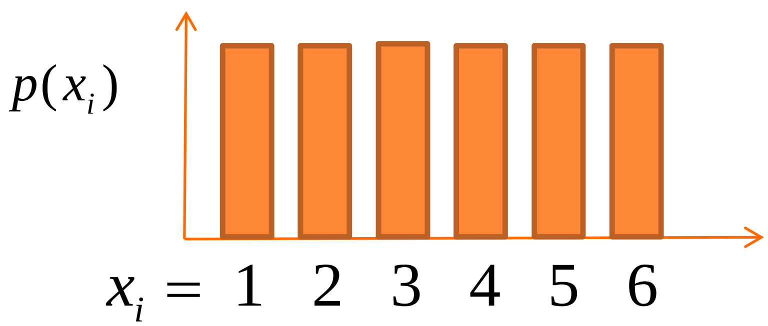
- Discrete Random Variable
  - X denotes a random variable
  - X can take on a countable number of values

$$X \in \{x_1, \ldots, x_n\}$$

  - The probability that X takes on a specific value

$$p(X = x_i) \text{ or } p(x_i)$$

  - A 6-sided die's discrete probability distribution

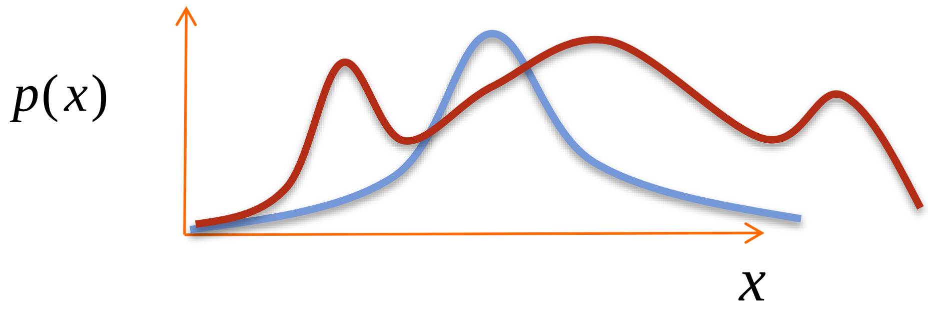


$$\sum_{i=1}^{n} p(x_i) = 1$$

PROBABILITY

- Continuous Random Variable
  - X takes on a value in a continuum
  - Probability density function, $p(X=x)$ or $p(x)$
  - Evaluated over finite intervals of the continuum

$$p(x \in (a,b)) = \int_a^b p(x)dx$$



PROBABILITY

- Measures of Distributions
  - Mean
    - Expected value of a random variable

$$\mu = E[X]$$

$$\mu = \sum_{i=1}^{n} x_i p(x_i) \text{ discrete case}$$

$$\mu = \int xp(x)dx \quad \text{continuous case}$$

  - Variance
    - Measure of the variability of a random variable

$$Var(X) = E\left[(X-\mu)^2\right]$$

$$Var(X) = \sum_{i=1}^{n} (x_i - \mu)^2 p(x_i) \text{ discrete case}$$

$$Var(X) = \int (x-\mu)^2 p(x)dx \quad \text{continuous case}$$

    - Square root of variance is standard deviation, $\sigma^2 = Var(X)$

Probability

- Multi-variable distributions
  - Vector of random variables

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

  - Mean

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_n] \end{bmatrix}$$

PROBABILITY

- Multi-variable distributions
  - Covariance
    - Measure of how much two random variables change together

$$Cov(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)]$$
$$= E[X_i X_j] - \mu_i \mu_j$$

    - If $Cov(X,Y)>0$, when X is above its expected value, then Y tends to be above its expected value
    - If $Cov(X,Y)<0$, when X is above its expected value, then Y tends to be below its expected value
    - If X, Y are independent, $Cov(X,Y) = 0$

PROBABILITY

- Multi-variable distribution
  - Covariance Matrix, Σ
    - Defines variational relationship between each pair of random variables

$$\Sigma_{i,j} = Cov(X_i, X_j)$$

    - Generalization of variance, diagonal elements represent variance of each random variable

$$Cov(X_i, X_i) = Var(X_i)$$

    - Covariance matrix is symmetric, positive semi-definite

PROBABILITY

- Multiplication by a constant matrix yields

$$
\begin{aligned}
\operatorname{cov}(Ax) &= E[(Ax - A\mu)(Ax - A\mu)^T] \\
&= E[A(x-\mu)(x-\mu)^T A^T] \\
&= AE[(x-\mu)(x-\mu)^T]A^T \\
&= A\operatorname{cov}(x)A^T
\end{aligned}
$$

PROBABILITY

- Addition/Subtraction of random variables

$$
\begin{aligned}
\operatorname{cov}(X \pm Y) &= E\left[\Big((X-\mu_x)\pm(Y-\mu_y)\Big)\Big((X-\mu_x)\pm(Y-\mu_y)\Big)^T\right] \\
&= E\Big[(X-\mu_x)(X-\mu_x)^T \pm (X-\mu_x)(Y-\mu_y)^T \\
&\qquad \pm (Y-\mu_y)(X-\mu_x)^T + (Y-\mu_y)(Y-\mu_y)^T\Big] \\
&= \operatorname{cov}(X) + \operatorname{cov}(Y) \pm 2\operatorname{cov}(X,Y)
\end{aligned}
$$

  - If X,Y independent,

$$\operatorname{cov}(X \pm Y) = \operatorname{cov}(X) + \operatorname{cov}(Y)$$

Probability

- Joint Probability
  - Probability of x and y:

$$p(X = x \text{ and } Y = y) = p(x, y)$$

    - e.g. probability of clouds and rain today

- Independence
  - If X, Y are independent, then

$$p(x, y) = p(x)p(y)$$

    - e.g. probability of two heads coin-flips in a row is ¼

PROBABILITY

- Conditional Probability
  - Probability of x given y

$$p(X = x \mid Y = y) = p(x \mid y)$$

    - Probability of KD for dinner, given a Waterloo engineer is cooking

  - Relation to joint probability

$$p(x \mid y) = \frac{p(x, y)}{p(y)}$$

  - If X and Y are independent,

$$p(x \mid y) = p(x)$$

    - Follows from the above

PROBABILITY

- Law of Total Probability

Discrete	Continuous
$\sum_{x} p(x) = 1$	$\int p(x)dx = 1$
$p(x) = \sum_{y} p(x,y)$	$p(x) = \int p(x,y)dy$
$p(x) = \sum_{y} p(x \mid y)p(y)$	$p(x) = \int p(x \mid y)p(y)dy$

PROBABILITY

- Probability distribution
  - It is possible to define a discrete probability distribution as a column vector

$$p(X = x) = \begin{bmatrix} p(X = x_1) \\ \vdots \\ p(X = x_n) \end{bmatrix}$$

  - The conditional probability can then be a matrix

$$p(x \mid y) = \begin{bmatrix} p(x_1 \mid y_1) & \dots & p(x_1 \mid y_m) \\ \vdots & \ddots & \vdots \\ p(x_n \mid y_1) & \cdots & p(x_n \mid y_m) \end{bmatrix}$$

PROBABILITY

- Discrete Random Variable
  - And the Law of Total Probabilities becomes

$$p(x) = \sum_y p(x \mid y)\, p(y)$$
$$= p(x \mid y) \cdot p(y)$$

  - Note, each column of $p(x \mid y)$ must sum to 1

$$\sum_x p(x \mid y) = \sum_x \frac{p(x,y)}{p(y)}$$

Relation of joint and conditional probabilities

$$= \frac{\sum_x p(y,x)}{p(y)} = \frac{p(y)}{p(y)} = 1$$

Total probability

PROBABILITY

- Bayes Theorem
  - From definition of conditional probability

$$p(x \mid y) = \frac{p(x,y)}{p(y)}, \quad p(y \mid x) = \frac{p(x,y)}{p(x)}$$

$$p(x \mid y)\,p(y) = p(x,y) = p(y \mid x)\,p(x)$$

  - Bayes Theorem defines how to update one's beliefs about X given a known (new) value of y

$$\boxed{p(x \mid y) = \frac{p(y \mid x)\,p(x)}{p(y)} = \frac{likelihood \cdot prior}{evidence}}$$

Probability

- Bayes Theorem
  - If Y is a measurement and X is the current vehicle state, Bayes Theorem can be used to update the state estimate given a new measurement
    - Prior: probabilities that the vehicle is in any of the possible states
    - Likelihood: probability of getting the measurement that occurred given every possible state is the true state
    - Evidence: probability of getting the specific measurement recorded

$$\boxed{p(x \mid y) = \frac{p(y \mid x)\, p(x)}{p(y)} = \frac{likelihood \cdot prior}{evidence}}$$

PROBABILITY

- Bayes Theorem
  - Example: Drug testing
    - A drug test is 99% sensitive (will correctly identify a drug user 99% of the time)
    - The drug test is also 99% specific (will correctly identify a non-drug user 99% of the time
    - A company tests its employees, 0.5% of whom are drug users
    - What's the probability that a positive test result indicates an actual drug user?

33%, 66%, 97%, 99%???

PROBABILITY

- Bayes Theorem
  - Example: Drug Testing
    - Employees are either users or non-users

$$X = \{u, n\}$$

    - The test is either positive or negative

$$Y = \{\rho, \eta\}$$

    - We want to find the probability that an employee is a user given the test is positive. Applying Bayes Theorem:

$$p(u \mid \rho) = \frac{p(\rho \mid u)\, p(u)}{p(\rho)} = \frac{likelihood \cdot prior}{evidence}$$

Probability

- Bayes Theorem
  - Example: Drug Testing
    - Prior: Probability that an individual is a drug user

$$p(u) = 0.005$$

    - Likelihood: Probability that a test is positive given an individual is a drug user

$$p(\rho \mid u) = 0.99$$

    - Evidence: Total probability of a positive test result

$$p(\rho) = p(\rho \mid u)\, p(u) + p(\rho \mid n)\, p(n)$$
$$= (0.99)(0.005) + (0.01)(0.995) = 0.0143$$

PROBABILITY

- Bayes Theorem
  - Example: Drug Testing
    - Finally, the probability an individual is a drug user given a test is positive

$$p(u \mid \rho) = \frac{p(\rho \mid u)\, p(u)}{p(\rho)}$$

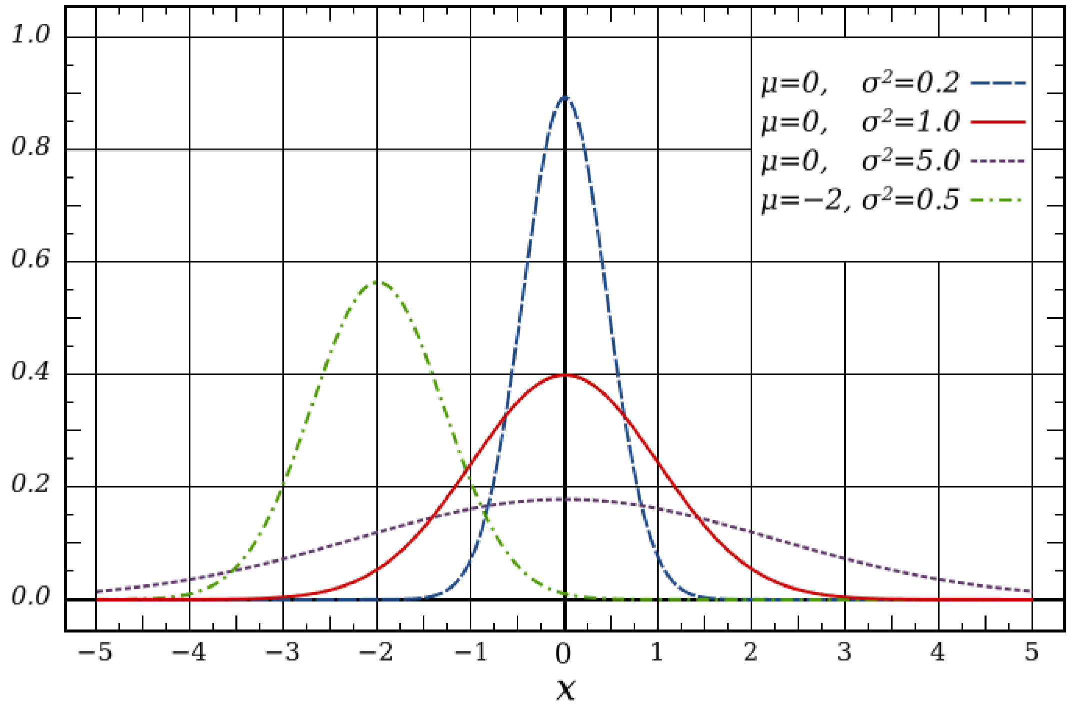
$$= \frac{(0.99)(0.005)}{(0.0149)} = 0.3322$$

    - 33% chance of that positive test result has caught a drug user. That's not a great test!
    - Difficulty lies in the large number of non-drug users that are tested
      - Hard to find a needle in the haystack with a low resolution camera.

Probability

- Gaussian Distribution (Normal)

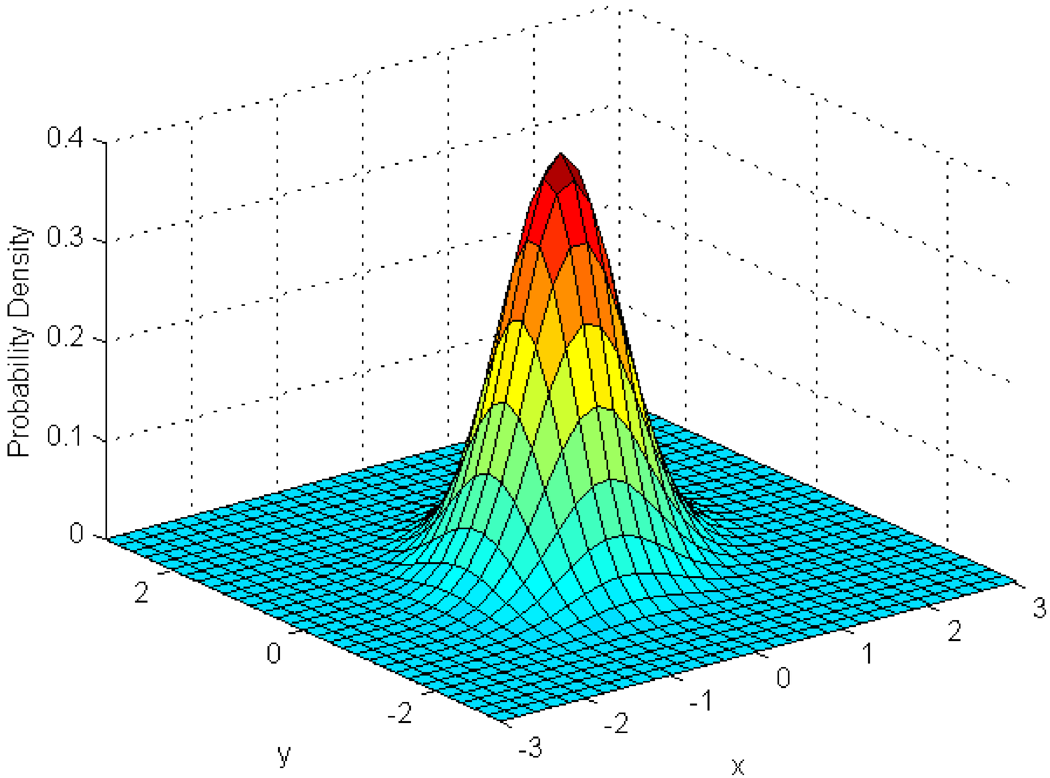
$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad p(x) \sim N(\mu, \sigma^2)$$



PROBABILITY

- Multivariate Gaussian Distribution (Normal)

$$p(\boldsymbol{x}) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{\mu})} \qquad p(\boldsymbol{x}) \sim N(\boldsymbol{\mu}, \Sigma)$$



PROBABILITY

- Properties of Gaussians
  - Linear combinations

$$x \sim N(\mu, \Sigma), \quad y = Ax + B$$

$$y \sim N(A\mu + B, A\Sigma A^T)$$

  - The result remains Gaussian!
    - Note: exclamation point, because this is somewhat surprising, and does not hold for multiplication, division.
      - Let's take a look

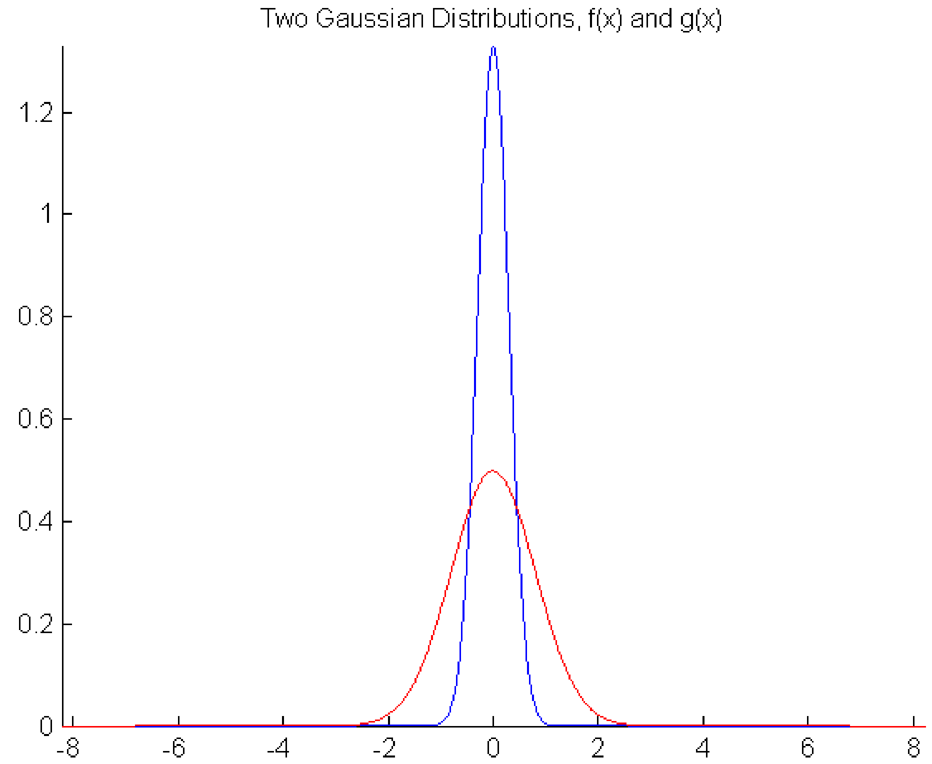
PROBABILITY

- Demonstration of combination of Gaussians
  - A tale of two univariate Gaussians
    - Define two Gaussians (zero mean)
    - Generate many samples from each distribution (5,000,000)
    - Combine these samples linearly, one sample from each distribution at a time
    - Multiply these samples
    - Divide these samples
    - Create histograms of the resulting samples
    - Take mean and variance of resulting samples
    - Generate Gaussian fit and compare

PROBABILITY

- Demonstration of combination of Gaussians
  - A tale of two univariate Gaussians

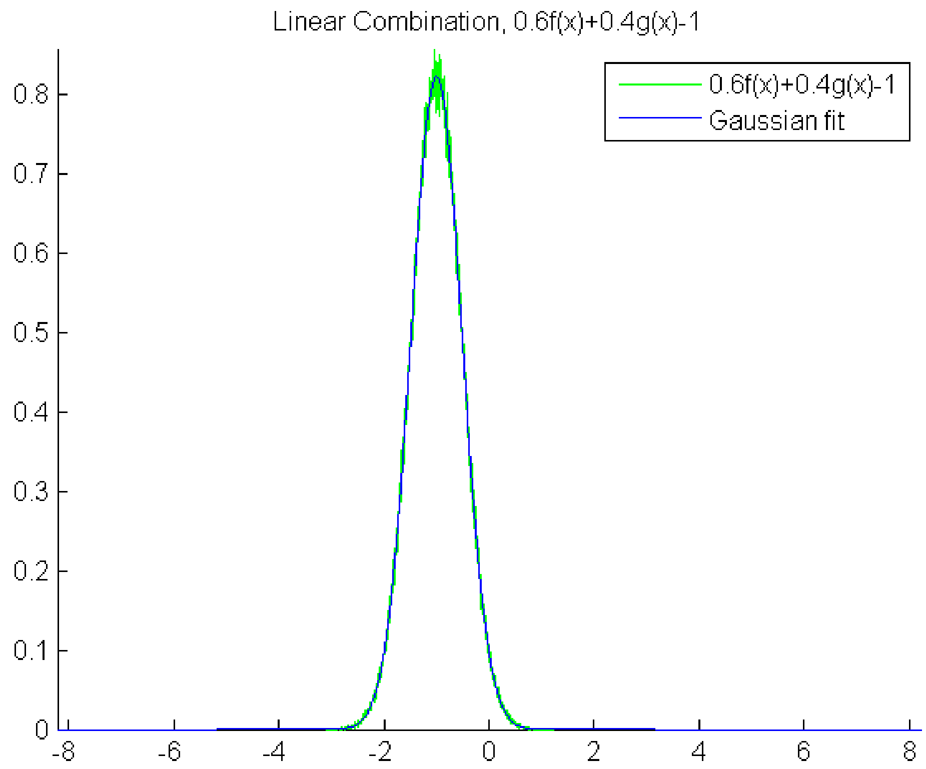
$$f(x) \sim N(0, 0.3^2) \qquad g(x) \sim N(0, 0.8^2)$$



PROBABILITY

- Demonstration of combination of Gaussians
  - Linear combination

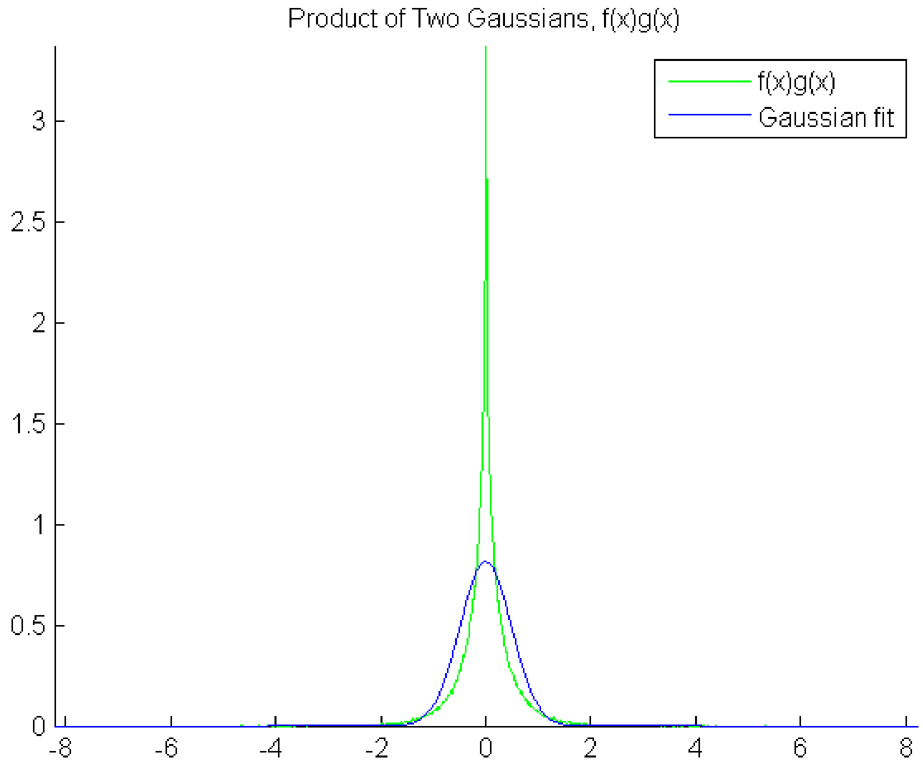
$$0.6f(x)+0.4g(x)-1 \sim N(-1,0.62^2)$$



PROBABILITY

- Demonstration of combination of Gaussians
  - Product

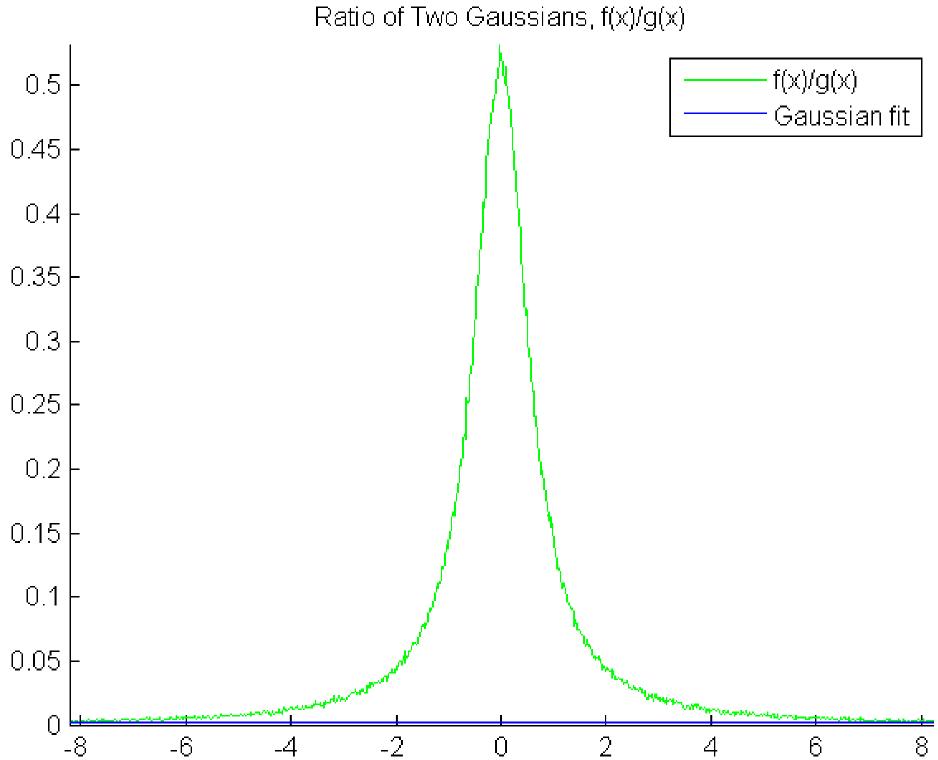
$$f(x)g(x) \sim N(0, 0.48^\wedge 2)$$



PROBABILITY

- Demonstration of combination of Gaussians
  - Quotient

$$f(x)/g(x) \sim N(0,170^\wedge 2)$$



Probability

- Generating multivariate random noise samples
  - Define two distributions, the one of interest and the standard normal distribution

$$\delta \sim N(\mu, \Sigma) \qquad\qquad \omega \sim N(0, I)$$

  - If the covariance is full rank, it can be diagonalized
    - Symmetry implies positive semi-definiteness

$$\begin{aligned}\Sigma &= E\lambda E^T \\ &= E\lambda^{1/2} I \lambda^{1/2} E^T \\ &= HIH^T\end{aligned}$$

  - Can now relate the two distributions (linear identity)

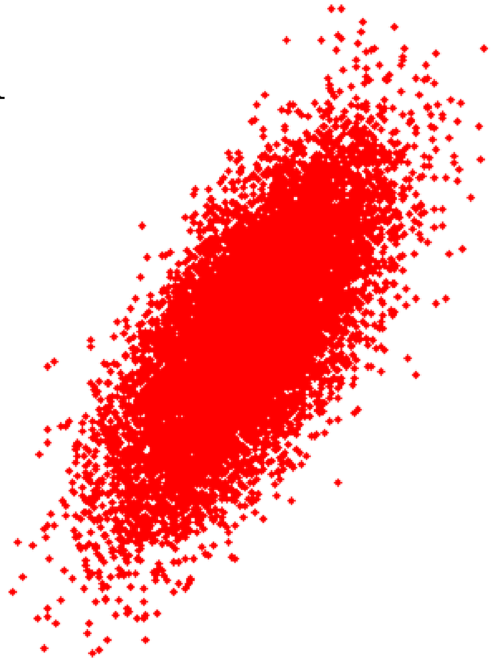
$$\begin{aligned}&\delta \sim N(\mu, HIH^T) \\ &\delta = \mu + H\omega\end{aligned}$$

Probability

- To implement this in Matlab for simulation purposes
  - Define $\mu,\sum$
  - Find eigenvalues , λ, and eigenvectors, E of $\sum$
  - The noise can then be created with

$$\delta = \mu + E\lambda^{1/2}\text{randn(n,1)}$$

$$\Sigma = \begin{bmatrix} 4 & 4 \\ 4 & 8 \end{bmatrix}$$



PROBABILITY

- Confidence ellipses
  - Lines of constant probability
    - Found by setting pdf exponent to a constant
    - Principal axes are eigenvectors of covariance
    - Magnitudes depend on eigenvalues of covariance

$$\mu = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 4 & 4 \\ 4 & 8 \end{bmatrix}$$

50%, 99% error ellipses
Not easily computed,
code provided