BMEG 802 – Advanced Biomedical Experimental Design and Analysis

Probability

Joshua G. A. Cashaback, PhD

Lecture Objectives

- Point Probabilities
    a. axioms
    b. definitions (complement, mutually exclusive, joint, marginal, conditional)
    c. rules (addition, subtraction, multiplication)
- Continuous Probabilities
    a. normal distribution (univariate and multivariate)
    b. joint, marginal, conditional

POINT PROBABILITIES

- axioms
- definitions (complement, mutually exclusive, joint, marginal, conditional)
- rules (addition, subtraction, multiplication)

Notation

- $S =$ sample space (all possible outcomes)
- $p(A) =$ probability of event A
- $A \cup B =$ union of events A and B
- $A \cap B =$ intersection of events A and B
- $p(B|A) =$ probability of B given A
- $p(A')$ or $p(A^C)$ or $p(\bar{A}) =$ complement probability of $p(A)$

Axioms

- The probability of an event is a non-negative real number

$$P(E) \geq 0$$

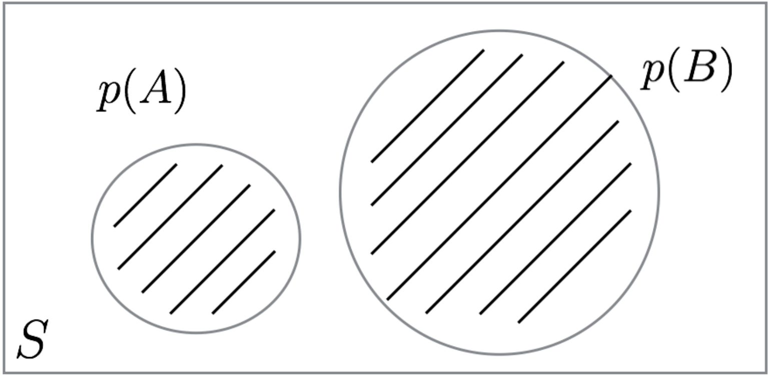
- The probability that at least one of the elementary events in the entire space will occur is 1

$$P(\Omega) = 1.$$

- Any countable sequence of disjoint sets (synonymous with mutually exclusive events) $E_1, E_2, \ldots$ satisfies

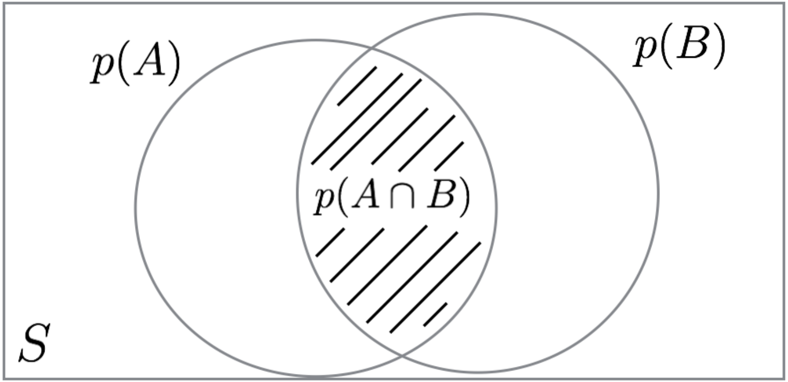
$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

Mutually Exclusive (Disjoint Probability)



$$p(A \cup B) = p(A) + p(B)$$
$$0.7 = 0.4 + 0.3$$

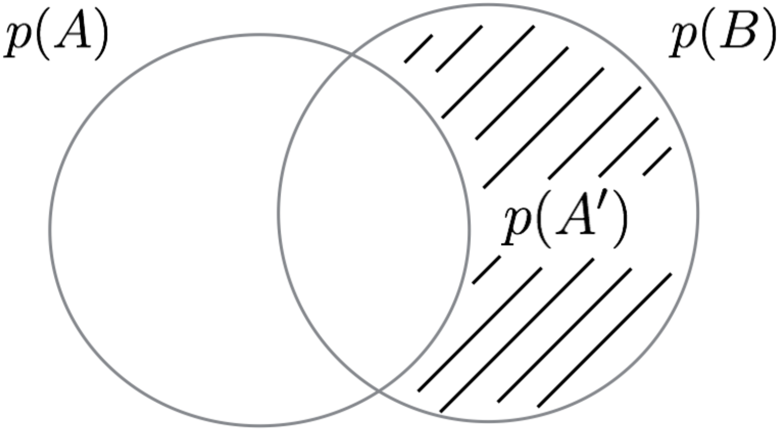
Joint Probability



$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$
$$0.5 = 0.4 + 0.3 - 0.2$$

- the probability of two events occuring simultaneously

Complement Probability



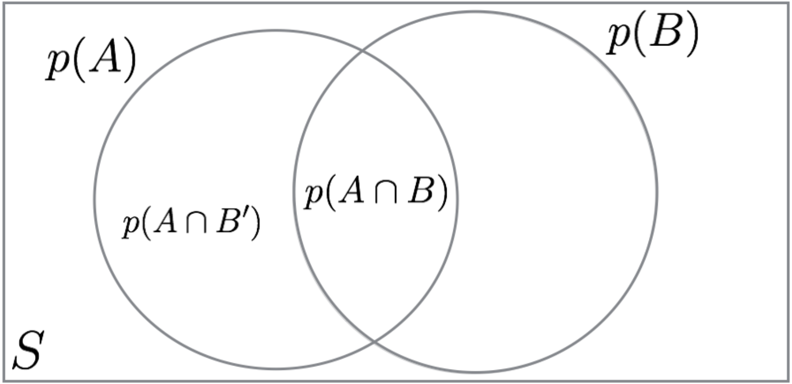
$$p(A') = 1 - p(A)$$

Marginal Probability

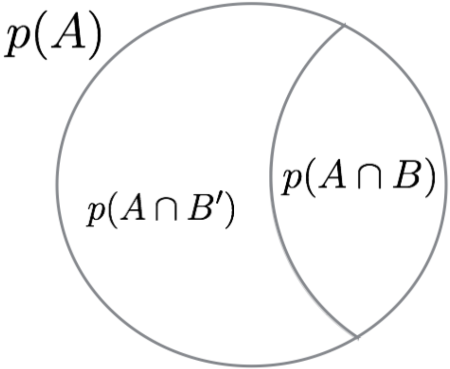
H \ L	Red	Yellow	Green	Marginal probability P(H)
Not Hit	0.198	0.09	0.14	**0.428**
Hit	0.002	0.01	0.56	**0.572**
Total	**0.2**	**0.1**	**0.7**	**1**

- Probability of a single event occurring (hit), independent of other events (light)
- e.g., probabilities of getting in an accident at an intersection irrespective of lights
- note: joint probabilities in each cell

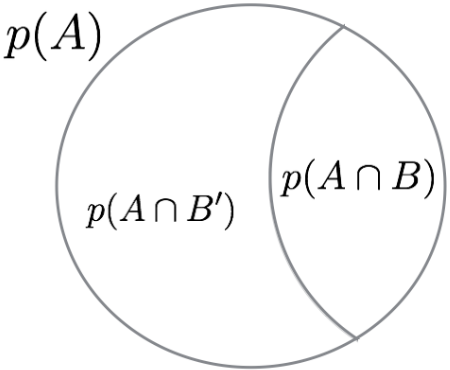
Marginal Probability



Marginal Probability

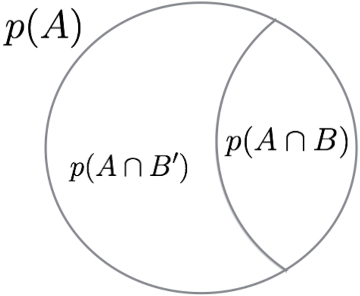


Marginal Probability



$$p(A \cap B') = p(A) - p(A \cap B)$$

Marginal Probability



$$p(A \cap B') = p(A) - p(A \cap B)$$

The marginal $p(A)$ or $p(B)$ is found by summating their disjoint parts.

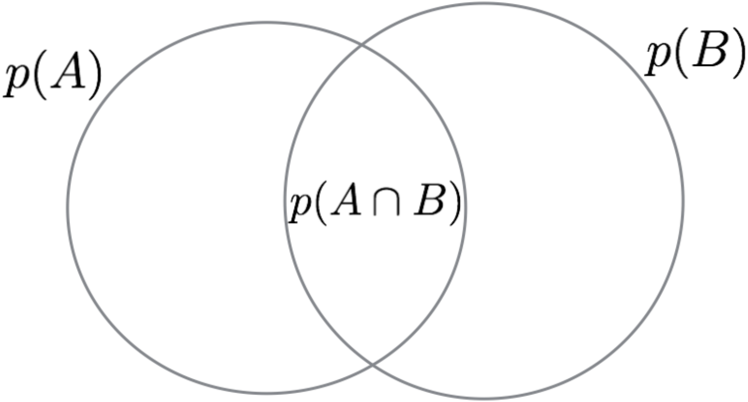
$p(A) = p(A \cap B) + p(A \cap B')$, and similarly

$$p(B) = p(A \cap B) + p(A' \cap B)$$

Conditional Probability

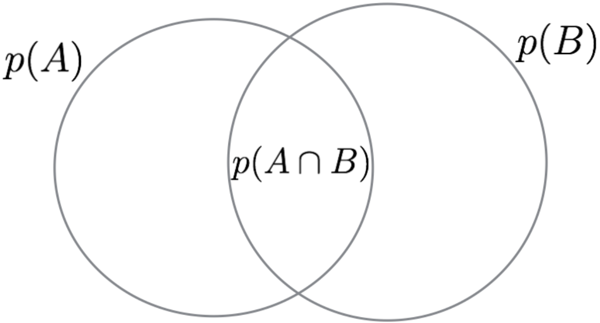
- $p(accepted) = 0.3$
- $p(funding|accepted) = 0.43$
- $p(funding \cap accepted) = p(funding|accepted) \cdot p(accepted)$
- $p(funding \cap accepted) = 0.43 \cdot 0.3 = 0.13$
- Probability that an event occurs given that another specific event *has already* occurred

Conditional Probability



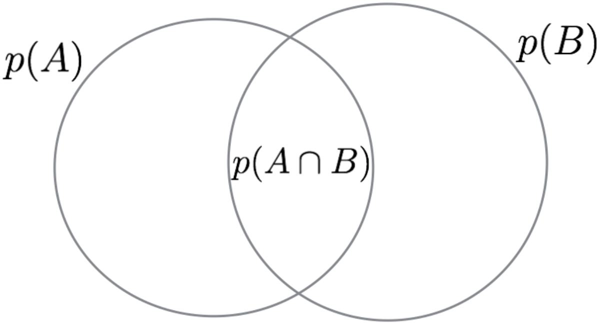
$$p(A \cap B) = p(B|A) \cdot p(A)$$
$$p(B|A) = \frac{p(A \cap B)}{p(A)}$$

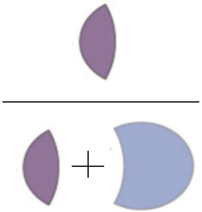
Conditional Probability



$$p(A \cap B) = p(B|A) \cdot p(A)$$
$$p(B|A) = \frac{p(A \cap B)}{p(A)}$$
$$p(A \cap B) = p(A|B) \cdot p(B) \text{ (in terms of B)}$$
$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

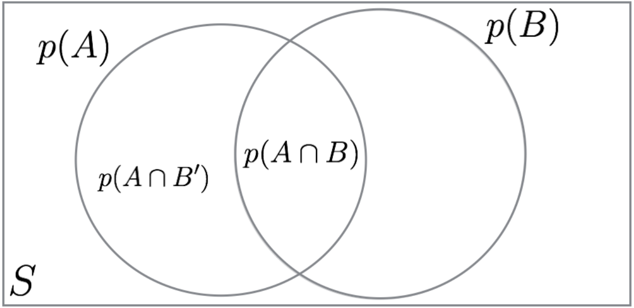
Conditional Probability



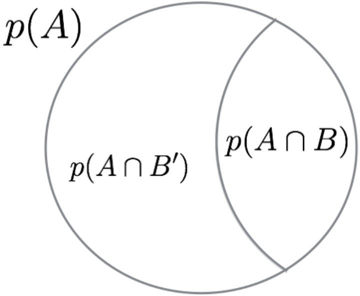


$$p(A \cap B) = p(B|A) \cdot p(A)$$
$$p(B|A) = \frac{p(A \cap B)}{p(A)}$$
$$p(A \cap B) = p(A|B) \cdot p(B) \text{ (in terms of B)}$$
$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

Conditional Probability Complements

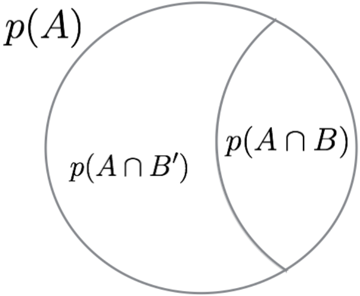


Conditional Probability Complements



$$p(A \cap B') = p(B'|A) \cdot p(A)$$

Conditional Probability Complements



$$p(A \cap B') = p(B'|A) \cdot p(A)$$

Other friendly complements:

$$p(A' \cap B) = p(B|A') \cdot p(A')$$

$$p(A' \cap B') = p(B'|A') \cdot p(A')$$

Probability Rules

- Rule of Subtraction: The probability that A will occur is equal to 1 minus the probability that A will NOT occur.

$$p(A) = 1 - P(A')$$

- Rule of Multiplication: The probability that Events A and B both occur is equal to the probability that Event A occurs times the probability that Event B occurs, given that A has occurred.

$$p(A \cap B) = p(A) \cdot p(B|A)$$

- Rule of Addition: The probability that Event A or Event B occurs is equal to the probability that Event B occurs minus the probability that both Events A and B occur.

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

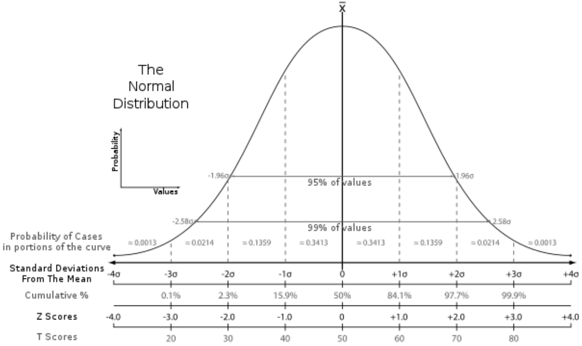
Note: we can redefine the addition rule given that $p(A \cap B) = p(A) \cdot p(B|A)$, such that:

$$p(A \cup B) = p(A) + p(B) - p(A) \cdot p(B|A)$$

CONTINUOUS PROBABILITIES

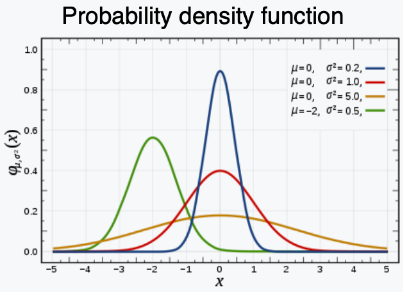
- Identical concepts to point probabilities!
- Normal distribution (univariate and multivariate)
- joint, marginal, conditional

Normal Distribution (Univariate)



- The probability that some value of x will occur
- Think of a histogram

Normal Distribution - Probability Density Function

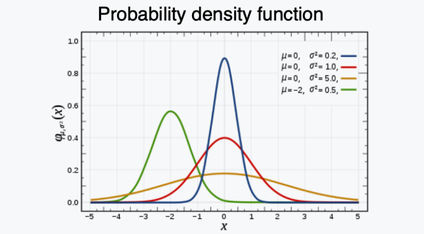


The red curve is the *standard normal distribution*

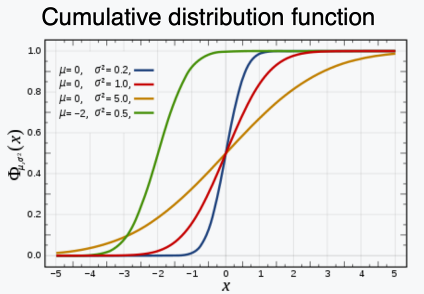
$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$\mathcal{N}(\mu,\sigma)$$

Normal Distribution - Cumulative Density Function



The red curve is the *standard normal distribution*



$$\Phi(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

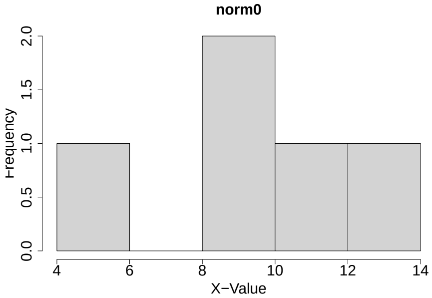
See Probability and Statistics Primer

Sampling from a Normal Distribution ($N = 5$)

```r
norm0 <- rnorm(5, mean=10, sd=2)
norm0
```

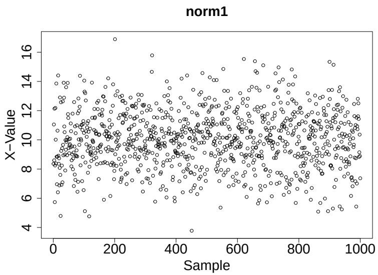
```
## [1]  8.331721  5.728901 10.472575 12.693787  8.883932
```

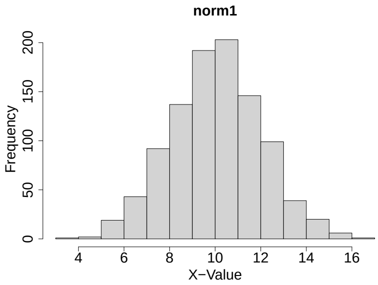
```r
hist(norm0, main="norm0", xlab="X-Value", ylab="Frequency",
     cex.lab=2.0, cex.axis=2.0, cex.main=2.0)
```



Sampling from a Normal Distribution (N = 1000)

```
norm1 <- rnorm(1000, mean=10, sd=2)
par(mar=c(5,5,5,5))
plot(norm1, main="norm1", xlab="Sample",
     ylab="X-Value", cex.lab=2.0, cex.axis=2.0, cex.main=2.0)
hist(norm1, main="norm1", xlab="X-Value", ylab="Frequency",
     cex.lab=2.0, cex.axis=2.0, cex.main=2.0)
```



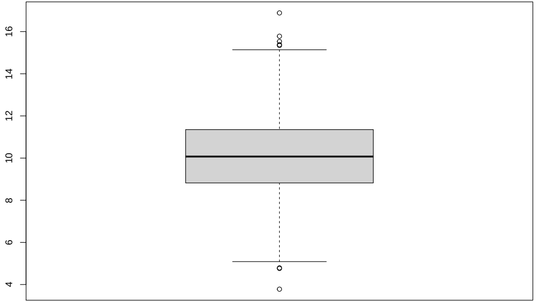


Summary of Sampled Data

```r
summary(norm1)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   3.780   8.822  10.073  10.056  11.351  16.888
```

```r
boxplot(norm1)
```



Fit the Data

install.packages("fitdistrplus")

```r
library(fitdistrplus)
```

```
## Loading required package: MASS
```

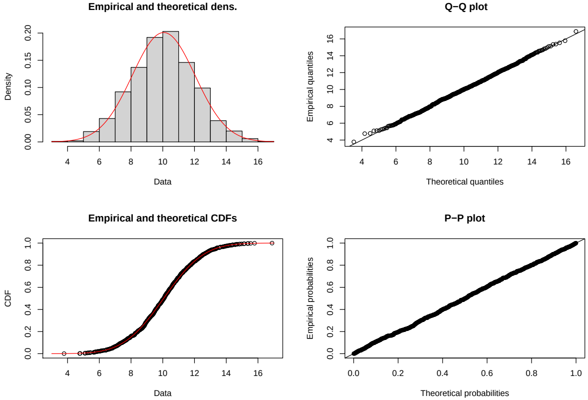
```
## Loading required package: survival
```

```r
FIT <- fitdist(norm1, "norm")
FIT
```

```
## Fitting of the distribution ' norm ' by maximum likelihood
## Parameters:
##        estimate Std. Error
## mean 10.056243 0.06278478
## sd    1.985429 0.04439549
```

Fit the Data

```r
plot(FIT)
```



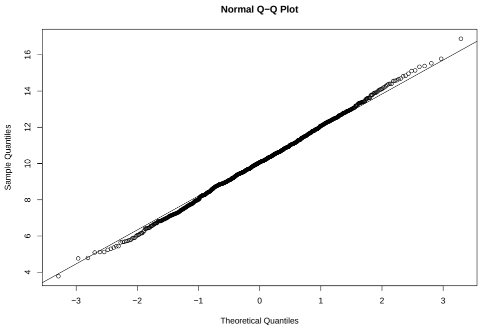
Q-Q (Quantile-Quantile) compares quantiles (divide data into n parts), P-P(Probability-Probability) compares CDF. Both test 'normality'

Testing Normality - QQ plot

```
library("car")
```

```
## Loading required package: carData
```

```
qqnorm(norm1)
qqline(norm1)
```



Testing Normality - Shapiro-Wilk Test (Univariate)

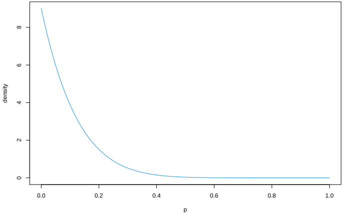
```
library("car")
shapiro.test(norm1)
```

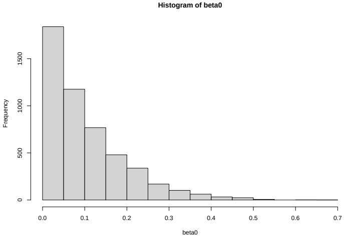
```
##
##  Shapiro-Wilk normality test
##
## data:  norm1
## W = 0.99921, p-value = 0.9586
```

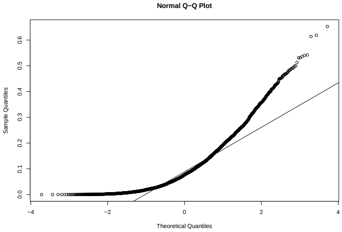
p-value $> 0.05 =$ normally distributed (we'll get more into p-values next class)

Testing Normality - Beta Distribution

```r
beta0 <- rbeta(5000, 1, 9)
p = seq(0,1, length=100)
plot(p, dbeta(p, 1, 9), ylab="density", type ="l", col=4)
hist(beta0)
qqnorm(beta0)
qqline(beta0)
```







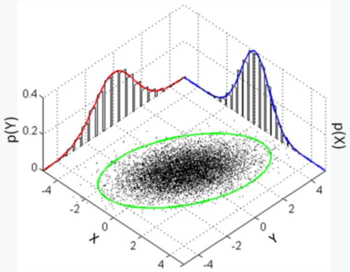
Testing Normality - Beta Distribution

```r
shapiro.test(beta0)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  beta0
## W = 0.86643, p-value < 2.2e-16
```

Normal Distribution (Multivariate)

Probability density function



Many sample points from a multivariate normal distribution with $\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 3/5 \\ 3/5 & 2 \end{bmatrix}$, shown along with the 3-sigma ellipse, the two marginal distributions, and the two 1-d histograms.

e.g., $X = \text{height}$, $Y = \text{weight}$

Normal Distribution (Bivariate PDF)

$$f_{\mathbf{X}}(x_1,\ldots,x_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k \mid \boldsymbol{\Sigma}|}} \qquad (\textit{bivariate}, k=2)$$

$$\boldsymbol{\mu} = \begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix}\sigma_X^2 & r\sigma_X\sigma_Y\\ r\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}.$$

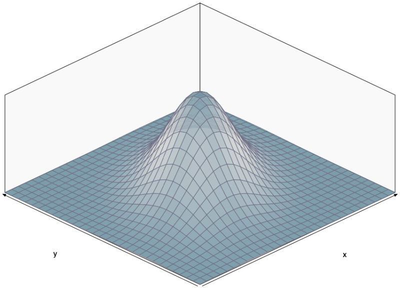
OR

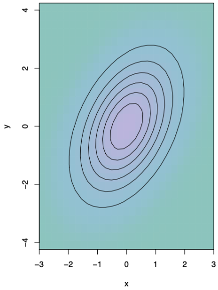
$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[(\frac{x-\mu_X}{\sigma_X})^2-2r(\frac{x-\mu_X}{\sigma_X})(\frac{y-\mu_Y}{\sigma_Y})+(\frac{y-\mu_Y}{\sigma_Y})^2\right]}$$

r = correlation between X and Y

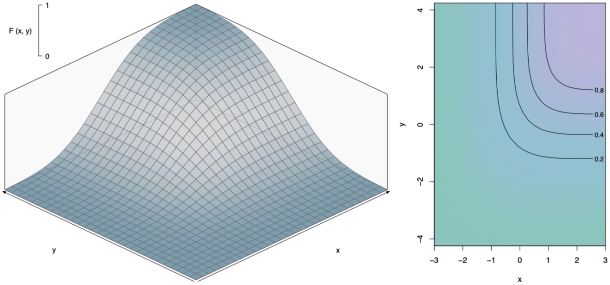
Σ is positive definite

Normal Distribution (Bivariate PDF)





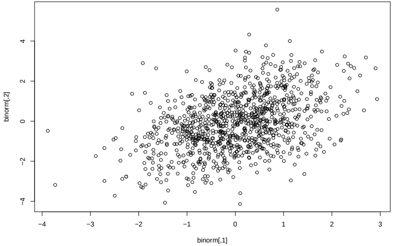
Normal Distribution (Bivariate CDF)



Sample from a Bivariate Normal

install.packages("MASS")

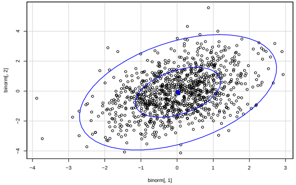
```
library("MASS")
my_n1 <- 1000
my_mu1 <- c(0, 0)
my_Sigma1 <- matrix(c(1, 0.6, 0.6, 2), ncol = 2)
binorm <- mvrnorm(n = my_n1, mu = my_mu1, Sigma = my_Sigma1)
plot(binorm)
```



Confidence Ellipse

install.packages("car")

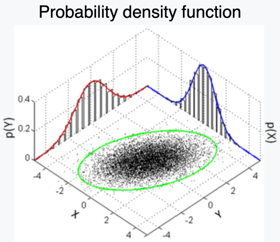
```r
library(car)
dataEllipse(binorm[,1], binorm[,2], levels=c(0.5, 0.975))
```



Eigendecomposition on Covariance Matrix

Square root of eigenvalues = principle axes

Joint Probability

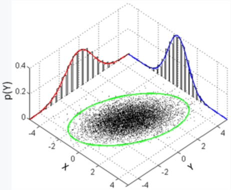


Many sample points from a multivariate normal distribution with $\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 3/5 \\ 3/5 & 2 \end{bmatrix}$, shown along with the 3-sigma ellipse, the two marginal distributions, and the two 1-d histograms.

- The bivariate normal distribution IS an example of a Joint Distribution
  - $p(X \cap Y)$
- We have already defined its PDF and CDF
- e.g., what is the probability you are x cm tall and weigh y kg?

Marginal Probability

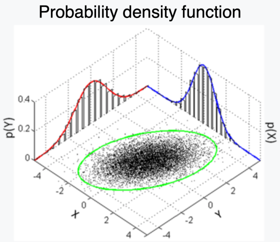
Probability density function



Many sample points from a multivariate normal distribution with $\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 3/5 \\ 3/5 & 2 \end{bmatrix}$, shown along with the 3-sigma ellipse, the two marginal distributions, and the two 1-d histograms.

- Represented as the univariate normal distributions on the ‘walls’
- Simply drop terms from the mean vector and covariance matrix related to the variable you want to marginalize out
- $p(X) = \mathcal{N}\left(\mu_X, \sigma_X^2\right)$ and $p(Y) = \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$
- e.g., what is the probability you weigh y kg?

Conditional Probability



Many sample points from a multivariate normal distribution with $\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 3/5 \\ 3/5 & 2 \end{bmatrix}$, shown along with the 3-sigma ellipse, the two marginal distributions, and the two 1-d histograms.

$$p(X \mid Y = y) = \mathcal{N}\left(\mu_X + \frac{\sigma_X}{\sigma_Y} r(y - \mu_Y),\ (1 - r^2)\sigma_X^2\right).$$

- e.g., probability that you are x cm tall given you weigh y kg
- $(1 - r^2)\sigma_X^2$ is the variance, so take square root to find SD