STAT 270 - Chapter 5
Continuous Distributions

June 27, 2012

Continuous rv's

Definition: X is a continuous rv if it takes values in an interval, i.e., range of X is continuous.
e.g. Temprature in degrees celcius in class.
Definition: **Probability density function** (pdf) or density of continuous a rv X, $f_X(x) \geq 0$ is such that:

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

for all $a < b$.
From definition:

$$P(X = c) = \int_c^c f_X(x)dx = 0$$

for all $c \in \Re$.

Properties of pdf

1. $f(x) \geq 0$ for all x
2. $\int f(x)dx = 1$ where the integral is taken over the range of X.

Example. The pdf of X is ,

$$f(x) = \begin{cases} 2x & 0 < x \leq \frac{1}{2} \\ \frac{3}{2} & \frac{1}{2} < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

cdf of a continuous rv

- Definition: is the same as the case of discrete rv's.
- Evaluation: needs integration.
- Consider continuous rv X with pdf $f(x)$,

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy$$

Percentile

100p-th percentile of a continuous distribution with cdf $F(x)$ is $\eta(p)$ such that

$$p = F(\eta(p)) = P(X \leq \eta(p))$$

i.e., 100p percent of the values fall below $\eta(p)$
e.g. median: 50-th percentile

Uniform distribution

$X \sim uniform(a, b)$

pdf:

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

cdf:

$$F(x) = \int_a^x f(x)dx = \int_a^x \frac{1}{b-a}dx = \frac{x-a}{b-a}$$

100p-th percentile:

$$p = F(\eta(p)) \Rightarrow \eta(p) = p(b-a) + a$$

median:

$$\tilde{x} = .5(b-a) + a = \frac{a+b}{2}$$

Example (5.1, 5.3)

$$f(x) = \begin{cases} x & 0 < x \leq 1 \\ \frac{1}{2} & 1 < x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) =$$

$$\eta(p) =$$

Expectation

Definition is the same as the case of discrete rv's
In claculation the sum is replaced by integral:
X continuous rv with pdf $f(x)$

$$E(X) = \int_x xf(x)dx$$

where the integral is taken over the range of X.

Example

Consider the pdf of the rv Y,

$$f(x) = \begin{cases} \frac{y}{25} & 0 \leq x \leq 5 \\ \frac{2}{5} - \frac{y}{25} & 5 \leq x \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

(a) Obtain the cdf of Y.
(b) Calculate the 100p-th percentile of Y.
(c) Calculate $E(Y)$.

Example

Let X be a rv with the density function

$$f(x) = \begin{cases} \frac{x}{2} & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Calculate $P(X \leq 1)$.
(b) Calculate $P(0.5 \leq X \leq 1.5)$.
(c) Calculate $P(0.5 < X)$.

Example

Suppose I never finish the lectures before the end of the hour and always finish within two minutes after the hour. Let X be the time that elapses between the end of the hour and the end of the lecture and suppose the pdf of X is

$$f(x) = \begin{cases} kx^2 & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Evalaute k.
(b) What is the probability that the lecture ends within one minute of the end of the hour?
(c) What is the probability that the lecture continues beyond the hour for between 60 and 90 seconds?
(d) What is the probability that the lecture continues for at least 90 seconds beyond the end of the hour?

Example

The cdf of a continuous rv X is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{4} & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

(a) Calculate $P(0.5 \leq X \leq 1)$.
(b) Calculate the median of X.
(c) Calculate the pdf of X.
(d) Calculate $E(X)$.

Expectation and variance

Expectation of a continuous rv X with pdf $f(x)$ is

$$\mu = E(X) = \int xf(x)dx$$

where the integral is taken over the range of X.

Expectation of a function of X, $g(X)$, is given by

$$E(g(X)) = \int g(x)f(x)dx$$

. Variance of a continuous rv X is given by

$$\sigma^2 = var(X) = E(X - E(X))^2 = \int (x - E(X))^2 f(x)dx$$

$$var(X) = E(X^2) - (E(X))^2$$

The Normal/Gaussian/bell-shaped distribution

- The most important distribution of all!
- Widely applicable to statistical science problems
- Mathematically elegant

Definition: $X \sim Normal(\mu, \sigma^2)$ if the pdf of X is given by

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

where $-\infty < x < \infty$, $-\infty < \mu < \infty$ and $\sigma > 0$.

Important notes on the Gaussian distribution

- Family of distributions indexed by parameters μ and σ^2.
- Symmetric about μ: $f(\mu + \delta) = f(\mu - \delta)$, for all δ.
- $f(x)$ decreases exponentially as $x \to -\infty$ and $x \to \infty$, but it never touches 0.
- $f(x)$ in intractable: $\int_a^b f(x)dx$ has to be approximated using numerical methods.

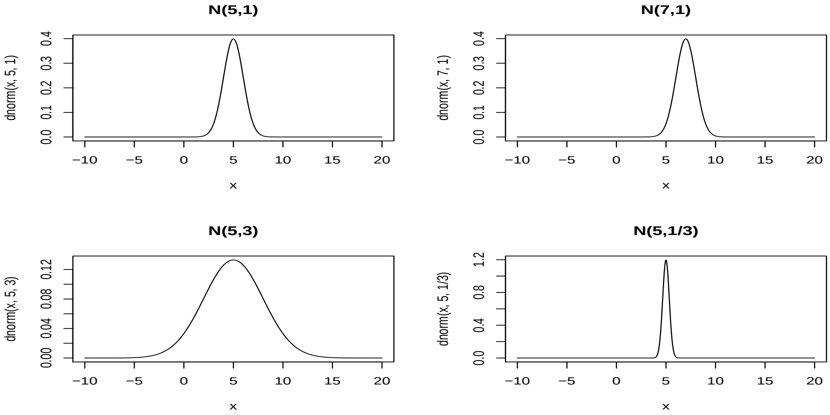
Mean and variance

If $X \sim Normal(\mu, \sigma^2)$ then

$$E(X) = \mu$$

and

$$var(X) = \sigma^2$$



The standard normal distribution

- Probabilities are obtained through normal tables
- Choice of μ and σ does not restrict us in calculation of probabilities: any normal distribution can be converted into the **standard** normal distribution: $Z \sim Normal(0, 1)$,

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

- Normal tables provide cumulative probabilities for the standard normal variable.

Some examples

- $P(Z \leq 2.04) = P(Z < 2.04) =$
- $P(Z > 2.08) = 1 - P(Z \leq 2.08) =$
- Symmetry: $P(Z > -1) = P(Z \leq 1) =$
- Interpolation (rarely needed):
  $P(Z < 2.03) \approx P(Z < 2.00) + \frac{3}{4}[P(Z < 2.04) - P(Z < 2.00)] =$
- Inverse problem: find z such that 30.5% of the standard normal population exceed z, i.e., $P(Z > z) = 0.305$.

Note: Little probability beyond ± 3.

Some useful Z-values

z	1.282	1.645	1.96	2.326	2.576
f(z)	0.9	0.95	0.975	0.99	0.995

Transformation to the standard normal distribution:

If

$$X \sim Normal(\mu, \sigma^2)$$

then

$$Z = \frac{X - \mu}{\sigma} \sim Normal(0, 1)$$

Example

A subset of Canadians watch an average of 6 hours of TV every day. If the viewing times are normally distributed with sd of 2.5 hours. What is the probability that a randomly selected person from thet population watches more than 8 hours of TV per day?

Example

The substrate concentration (mg/cm^3) of influent to a reactor is normally distrbuted with $\mu = 0.4$ and $\sigma = 0.05$.
(a) What is the probability that the concentration exceeds 0.35?
(b) What is the probability that the concentration is at most 0.2?
(c) How would you characterize the largest 5% of all concentration values?

More examples ...

Percentiles of the normal distribution

$\eta(p) =$ 100p-th percentile for $X \sim Normal(\mu, \sigma^2)$
$\eta_Z(p) =$ 100p-th percentile for $Z \sim Normal(0, 1)$

then

$$\eta(p) = \mu + \sigma \eta_Z(p)$$

because

Example

Find the 86.43th percentile of the Normal(3,25).

The normal approximation to the binomial distribution

$$X \sim binomial(n, p)$$

where p is neither too small nor too large
(as a rule of thumb, $np \geq 5$ and $n(1-p) \geq 5$)
then the distribution of X is "close" to distribution of Y where

$$Y \sim Normal(np, np(1-p))$$

Note: $E(X) = E(Y)$ and $var(X) = var(Y)$.

- Not rigorous since **closeness in distribution** is not defined.
- Special case of the Central Limit Theorem (section 5.6).
- When n is small **continuity correction** is needed to improve the approximation (read page 86).

Example

Let $X \sim binomial(100, 0.5)$ and $Y \sim Normal(50 = np, 25 = np(1-p))$.

$P(X \geq 60) = \sum_{x=60}^{100} \binom{100}{x} (0.5)^x (0.5)^{(100-x)} = 0.017$

$P(Y \geq 60) = P(Z \geq \frac{60-50}{5}) = 1 - P(Z \leq 2) = 0.022$

The gamma distribution

If $X \sim \textit{gamma}(\alpha, \beta)$ the pdf of X is given by

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$$

where $x > 0$, $\alpha > 0$, $\beta > 0$, and $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-\frac{x}{\beta}} dx$.

- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$
- For positive integer α, $\Gamma(\alpha) = (\alpha - 1)!$

The gamma distribution

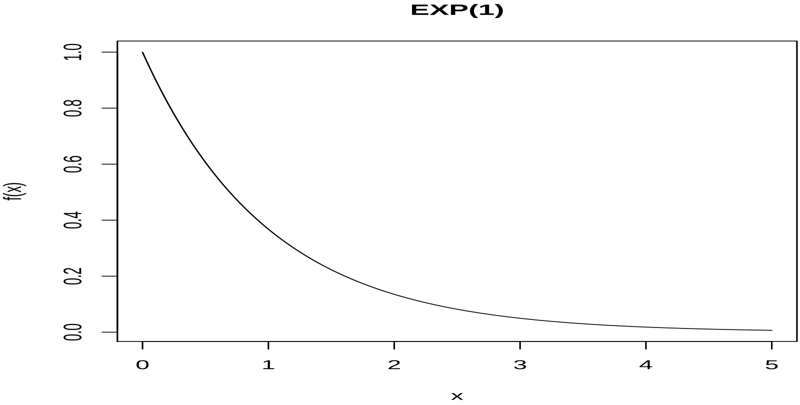
- Used to model right-skewed continuous data
- Family of distributions indexed by parameters α and β.
- Intractable except for particular values of α and β.
- $E(X) = \alpha\beta$ and $var(X) = \alpha\beta^2$
  Proof:

The exponential distribution

If $X \sim \textit{exponential}(\lambda)$ the pdf of X is given by

$$f(x) = \lambda e^{-\lambda x}$$

where $x > 0$ and $\lambda > 0$.



The exponential distribution

- Family of distributions indexed by the parameter λ.
- Special case of the gamma distribution where $\alpha = 1$ and $\beta = \frac{1}{\lambda}$. (1-parameter sub-family of the gamma family)
- $E(X) = \frac{1}{\lambda}$ and $var(X) = \frac{1}{\lambda^2}$
- cdf:

$$F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x}$$

  where $x > 0$.

Memoryless property

An old light bulb is just as good as a new one!

Let $X \sim \textit{exponential}(\lambda)$ be the life span in hours of a light bulb . What is the probability that it lasts $a + b$ hours given that it has lasted a hours?

$$P(X > a + b | X > a) = P(X > b)$$

Proof:

Do you believe this about light bulbs?

The relationship between the Poisson and exponential distributions

N_T: Number of events occurring in the time interval $(0, T)$

$$N_T \sim Poisson(\lambda T)$$

X: Waiting time until the first event occurs

cdf of X:

$$F(x) = 1 - e^{-\lambda x}$$

i.e.,

$$X \sim exponential(\lambda)$$

Example 5.14

Jointly distributed rv's

Multivariate data

- multiple measurements on subjects
- not always independent
- need to study the joint distribution of the variables to model multivariate data

Discrete rv: Joint pmf for $X_1, \ldots, X_m$

$$p(x_1, \ldots, x_m) = P(X_1 = x_1, \ldots, X_m = x_m)$$

Example 1

Suppose X and Y have a joint pmf given by the following table

	$X = 1$	$X = 2$	$X = 3$
$Y = 1$	.1	.5	.1
$Y = 2$	.05	.1	.15

$P(X = 3, Y = 2) =$
$P(X < 3, Y = 1) =$
Sum out the **nuisance** parameter X to obtain the **marginal** pmf of Y:
$p(Y = 1) =$
Verify that this is a joint pmf:

Read page 91.

Continuous rv's

Joint pdf of $X_1, \ldots, X_m$, $f(x_1, \ldots, x_m)$ is such that

- $f(x_1, \ldots, x_m) \geq 0$ for all $x_1, \ldots, x_m$.
- $\int \int \ldots \int f(x_1, \ldots, x_m) dx_1 dx_2 \ldots dx_m = 1$ where the integral is taken over the range of $X_1, \ldots, X_m$.

The probability of event A is given by

$$P((X_1, \ldots, X_m) \in A) = \int \int \ldots \int_A f(x_1, \ldots, x_m) dx_1 \ldots dx_m$$

Marginal pdf's are obtained by **integrating out** the **nuisance** variables.

Example 2

Let X and Y have the joint pdf $f(x,y) = \frac{2}{7}(x+2y)$ where $0 < x < 1$ and $1 < y < 2$.
(a) Calculate $P(X \leq \frac{1}{2}, Y \leq \frac{3}{2})$.
(b) Obtain $f(y)$.

Read Example 5.15.

Example 3 (dependence in range)

Let X and Y have the joint pdf

$$f(x) = \begin{cases} 3y & 0 \leq x \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Calculate $P(Y \leq \frac{2}{3})$.

Read Example 5.16.

Independent random variables

X and Y discrete independent rv's:

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

X and Y continuous rv's:

$$f(x, y) = f(x)f(y)$$

Example: Independent bivariate normal distribution

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\{-\frac{1}{2}(\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2})\} = f(x)f(y)$$

where $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$.

Read Example 5.18.

Conditional distributions

Conditional density or pdf

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Conditional pmf

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Example: Obtain $f_{X|Y=y}(x)$ for Example 2.

Read Example 5.19.

Expectation of functions of multiple rv's

Let $X_1, \ldots, X_m$ have joint pmf $p(x_1, \ldots, x_m)$ then

$$E(g(X_1, \ldots, X_m)) = \sum_{x_1} \ldots \sum_{x_m} g(x_1, \ldots, x_m) p(x_1, \ldots, x_m)$$

Similarly if $X_1, \ldots, X_m$ have joint pdf $f(x_1, \ldots, x_m)$ then

$$E(g(X_1, \ldots, X_m)) = \int \ldots \int g(x_1, \ldots, x_m) f(x_1, \ldots, x_m) dx_1 \ldots dx_m$$

Read Example 5.20.

Example 5.21

Suppose X and Y are independent with pdf's $f_X(x) = 3x^2$, $0 < x < 1$ and $f_Y(2y)$, $0 < y < 1$ respectively. Obtain $E(|X - Y|)$.

Covariance

$$cov(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

Proof of the last equality:

Correlation:

$$\rho = corr(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$$

- population analogue of r (sample correlation)
- describes the degree of linear relationship between X and Y.

Read Example 5.22.

Remark: If X and Y are independent rv's then

$$cov(X, Y) = corr(X, Y) = 0$$

Proof: page 97

Proposition: X and Y rv's,

$$E(a_1X + a_2Y + b) = a_1E(X) + a_2E(Y) + b$$

$$var(a_1X + a_2Y + b) = a_1^2 var(X) + a_2^2 var(Y) + 2a_1a_2cov(X, Y)$$

Proof: page 98

Generalization to m rv's $X_1, \ldots, X_m$,

$$E(\sum_{i=1}^{m} a_iX_i + b) = \sum_{i=1}^{m} a_iE(X_i) + b$$

$$var(\sum_{i=1}^{m} a_iX_i + b) = \sum_{i=1}^{m} a_i^2 var(X_i) + 2\sum_{i<j} a_ia_jcov(X_i, X_j)$$

Statistics

Definition: A statistic is a function of data, e.g., $\tilde{x}$, $\bar{x}$, s^2, $x_{(1)}$ etc.

- does not depend on unknown parameters
- $X_1, \ldots, X_n$ random $\Rightarrow Q(X_1, \ldots, X_n)$ random
  $x_1, \ldots, x_m$ a realization of $X_1, \ldots, X_n \Rightarrow Q(x_1, \ldots, x_n)$ a realization of $Q(X_1, \ldots, X_n)$, e.g., $\bar{x}$ is a realization of $\bar{X}$

Example (continued)

Obtain the distribution of $Q = |X - Y|$ where the joint pmf of X and Y is given in Example 1.
Solution:

Q	$X = 1$	$X = 2$	$X = 3$
$Y = 1$	0	1	2
$Y = 2$	1	0	1

Distribution of statistics

Discrete case,

$$p_Q(q) = \sum \cdots \sum_A p(x_1, \ldots, x_m)$$

Continuous case,

$$f_Q(q) = \int \cdots \int_A p(x_1, \ldots, x_m) dx_1 \ldots dx_m$$

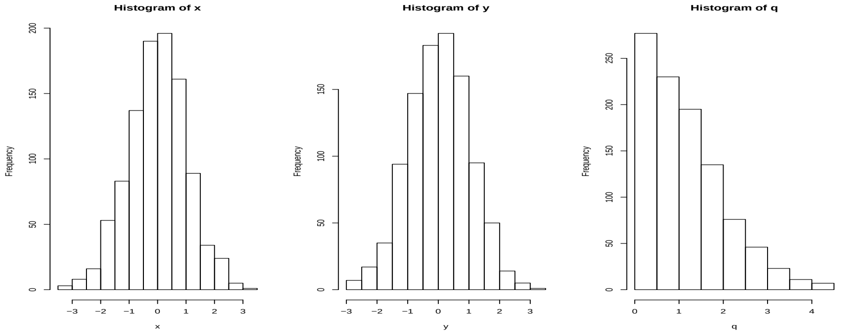
where $A = \{(x_1, \ldots, x_m) : Q(x_1, \ldots, x_m) = q\}$.

- Usually not easy to derive
- The distribution is studied using simulation

Simulation

1. Generate N copies of the sample $x_1, \ldots, x_n$ from its distribution
2. Evaluate $Q(x_1, \ldots, x_n)$ for each sample to get $q_1, \ldots, q_N$ which are N values generated from the distribution of Q.
3. Draw histograms, calculate summary statistics, etc.

Example: $X \sim N(0,1)$, $Y \sim N(0,1)$, $Q = |X - Y|$.



iid rv's

$X_1, \ldots, X_n$ independent and identically distributed (iid)

Random sample: A realization of $X_1, \ldots, X_n$: $x_1, \ldots, x_m$

Example: Let $X_1, \ldots, X_n$ be iid with $E(X_i) = \mu$ and $var(X_i) = \sigma^2$ then

$$E(\bar{X}) = \mu$$

and

$$var(\bar{X}) = \frac{\sigma^2}{n}$$

(less variation in $\bar{X}$ than in X_is).

$X_1, \ldots, X_n$ iid $N(\mu, \sigma^2) \Rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

Example

Linear combinations of normal random variables are normal random variables.

Let X, Y, Z be independent with distributions $N(1, \frac{1}{2})$, $N(0, \frac{3}{2})$ and $N(1, \frac{3}{2})$ respectively. What is the distribution of $W = 2X + Y - Z$?

Central Limit Theorem

- Most beautiful theorem in mathematics
- Few assumptions $\rightarrow$ important and useful results

Theorem (CLT): Let $X_1, \ldots, X_n$ be iid with $E(X_i) = \mu$ and $var(X_i) = \sigma^2$. Then as $n \rightarrow \infty$ the distribution of the statistic $Q(X) = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$ converges to a standard normal variable in distribution.

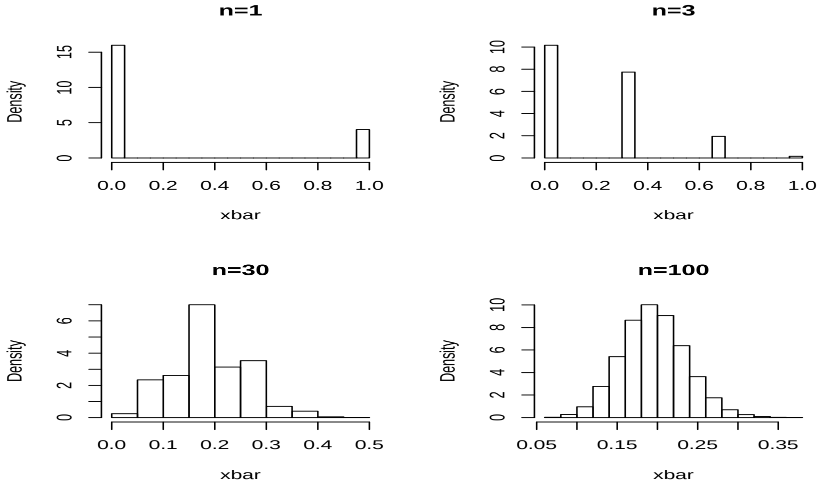
Note that:

- no assumptions are made for the distribution of X_is.
- Convergence in distribution is different from the common meaning of convergence in calculous.

Use $N \geq 30$ as a rule of thumb to apply CLT and conclude $\bar{X} \sim N(\mu, \sigma^2/n)$.

CLT

Example: Let X_i's, $i = 1, \ldots, n$, have a Bernoulli distribution with parameter $p = 0.2$. Want to study the distribution of $\bar{X}_n$ as n gets large.



Example

Suppose the human body weight average is $75kg$ with variance $400kg^2$. A hospital elevator has a maximum load of $3000kg$. If 40 people are taking the elevator what is the probability that the maximum load is exceeded?

Example

An Instructor gives a quiz with two parts. For a randomly selected student let X and Y be the scores obtained on the two parts respectively. The joint pmf of X and Y is given below:

$p(x,y)$	$y=0$	$y=5$	$y=10$	$y=15$
$x=0$	.02	.06	.02	.1
$x=5$	.04	.15	.2	.1
$x=10$	.01	.15	.14	.01

(a) What is the expected total score $E(X+Y)$?
(b) What is the expected maximum score from the two parts?
(c) Are x and Y independent?
(d) Obtain $P(Y=10|X\geq 5)$.

Example

Suppose $X_1 \sim N(1, .25)$ and $X_2 \sim N(2, 25)$ and $corr(X_1, X_2) = 0.8$. Obtain distribution of $Y = X_1 - X_2$.

Example

Suppose that the waiting time for a bus in the morning is uniformly distributed on $[0, 8]$ whereas the waiting time for a bus in the evening is uniformly distributed on $[0, 10]$. Assume that the waiting times are independent.

(a) If you take a bus each morning and evening for a week, what is the total expected waiting time?

(b) What is the variance of the total waiting time?

(c) What are the expected value and variance of how much longer you wait in the evening than in the morning on a given day?

Example

Tim has three errands where X_i is the time required for the ith errand, $i = 1, 2, 3$, and X_4 is the total walking time between errands. Suppose X_is are independent normal random variables with means $\mu_1 = 15$, $\mu_2 = 5$, $\mu_3 = 8$, $\mu_4 = 12$ and sd's $\sigma_1 = 4$, $\sigma_2 = 1$, $\sigma_3 = 2$, $\sigma_4 = 3$. If Tim plans to leave his office at 10am and post a note on the door reading "I will return by t am", what time t ensures that the probability of arriving later than t is .01?

Example

Let $X_1, \ldots, X_n$ be independent rv's with a uniform distribution on $[a, b]$.
Let $Y = max(X_1, \ldots, X_n)$. $E(Y) =?$

Example

Suppose that the bus 143 arrival times follow a poisson process with rate $\lambda = 5$ per hour. I arrive at the bus stop at 8:30 and meet one of my friends who tells me that she has already been waiting for the bus for 15 minutes. What is the probability that we take the bus no earlier than 8:45?