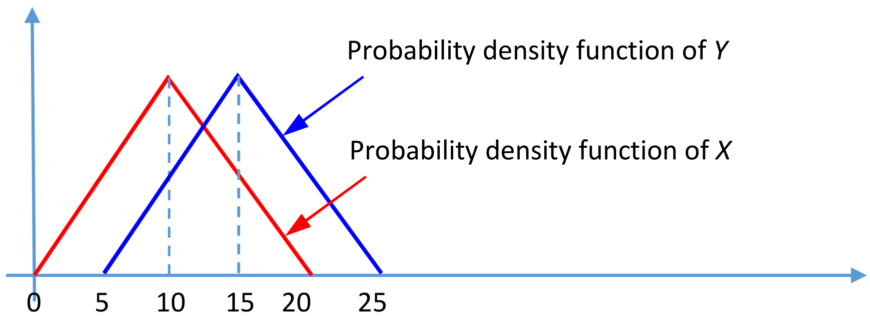
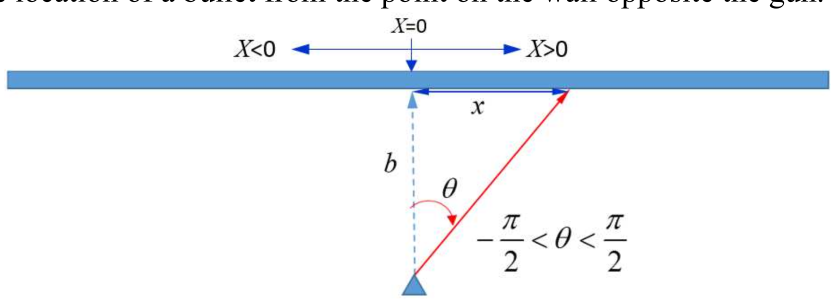
Statistics - Homework 3 (Due Nov. 8, 2017)

1. The probability density functions of two independent variables X and Y are shown in the following figure. If one random number (x or y) is drawn from each random variable, what is the probability that x is larger than y, i.e. $P[X > Y] = ?$



2. Suppose that a machine gun is mounted at a distance b from an infinitely long straight wall. If the angle of fire (θ) measured from the perpendicular from the gun to the wall is equally likely to be anywhere between $-\pi/2$ and $+\pi/2$. A random experiment is conducted by firing the machine gun. Let X represent the location of a bullet from the point on the wall opposite the gun.



(1) Derive the probability density function and cumulative probability function of X.

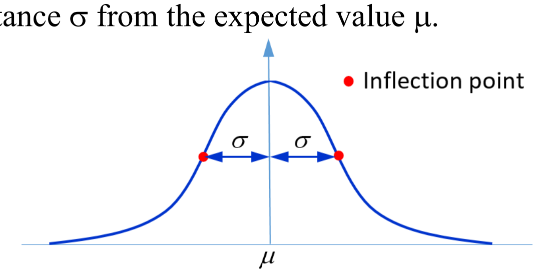
(2) Write an R code to simulate this random experiment. The number of firing is set to $N = 100000$ and b = 100 m. Plot the relative frequency histogram of X. [Note: Use X values falling in the range of (–5000, 5000) to plot the relative frequency histogram and set breaks=seq(–5000, 5000, 5) in the *hist* function of R. The relative frequency histogram of X in the range of (–5000, 5000) is NOT exactly the same as the relative frequency histogram of X, but should be very close.]

(3) Plot the probability density function of X on top of the above relative frequency histogram.

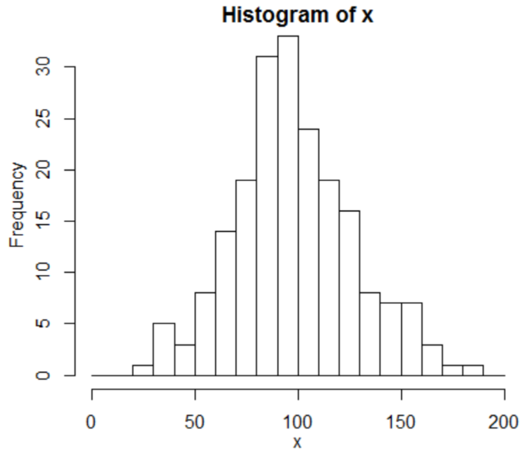
(4) In the same figure, plot the *empirical* cumulative distribution function and the CDF of X.

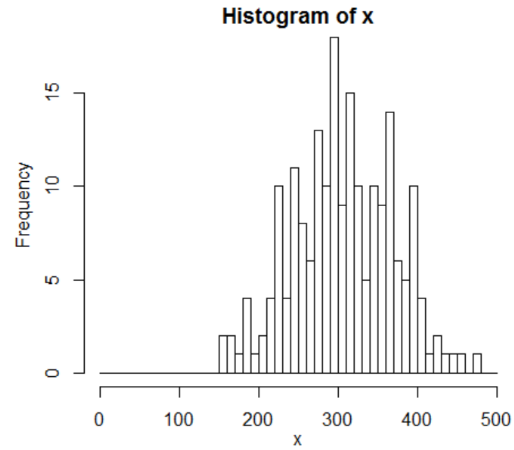
3. An inflection point is a place where a function switches from having a positive curvature (second derivative) to a negative curvature or vice-versa. In other words, it is a point where the second derivative equals zero.

(1) For a normal distribution $N(\mu,\ \sigma^2)$, show that the inflection point of its probability density function falls at a distance σ from the expected value μ.

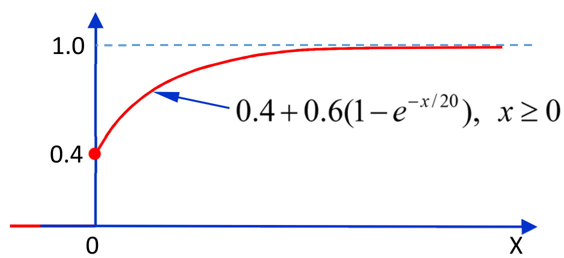


(2) Each of the following figures represents the histogram of 200 observed values from a normal distribution. Give you best guess of the standard deviation in each figure.





4. Let X be a normally distributed random variable, i.e., $X \sim N(\mu, \sigma)$.
   (1) Show that the kurtosis of X is 3.
   (2) Let $Y = (X-\mu)^2$. Find the variance of Y.
5. Prove that the kurtosis of any random variable cannot be less than 1.
6. A random variable X has the following cumulative distribution function.

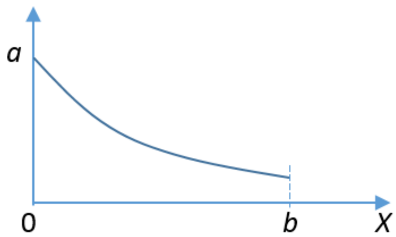


Calculate the expected value and standard deviation of X.

7. Let X_i, $i = 1,2,\cdots,n$, be independently and identically distributed exponential random variables with expected value λ. Let $Y = \sum_{i=1}^{n} X_i$. Prove that Y has a gamma distribution and find the probability density function of Y.
8. Let X_i, $i = 1,2,\cdots,n$, be independent Poisson random variables with expected value λ_i. Let $Y = \sum_{i=1}^{n} X_i$. Prove that Y has a Poisson distribution and find the probability density function of Y.
9. On average, there are 36 fatal traffic accidents in a city per year (365 days). Occurrence of fatal traffic accidents in any given day can be considered as rare.
   (1) Assuming a fatal accident just occurred, what is the probability that the next fifth fatal accident will occur no more than 30 days from now? Calculate this probability by considering the gamma distribution.
   (2) Calculate the same probability in (1) by considering the Poisson distribution.
10. Let X be a random variable with a continuous uniform density between -1 and 1, i.e., $X \sim U(-1,\ 1)$. Random variable Y is defined by the following transformation:

$$Y = \frac{2}{\pi}\sqrt{1 - X^2}$$

   (1) $Var(Y) = ?$
   (2) $P\left(Y \le \dfrac{1}{\pi}\right) = ?$
11. Let X be a random variable with probability density function $f_X(x) = ae^{-x} \ (0 \le x \le b,\ a > 0)$ where a is a constant.



(1) Find the first moment, second moment, and variance (E(X), E(X^2), Var(X)) of *X*.

(2) Find the cumulative distribution function of *X*.

(3) Generate a set of 10,000 random numbers of the above probability density with *a*=2 using the Acceptance/Rejection method and plot the histogram of the generated data.

(4) Calculate the sample mean and sample variance of the data generated in (3).