IS 604 Assignment 2

David Stern

September 19, 2015

1. Suppose that X is a discrete random variable having probability function $Pr(X = k) = ck^2$ for $k = 1, 2, 3$. Find c, $Pr(X \leq 2)$, E[X], and Var(X).

First, we will solve for c by summing the probabilities:

$$P(x = 1) = 1^2c$$
$$P(x = 2) = 2^2c$$
$$P(x = 3) = 3^2c$$

$$P(X = 1) + P(X = 2) + P(X = 3) = 1$$
$$c + 4c + 9c = 1$$
$$c = \frac{1}{14}$$

To find the probability that x is less than or equal to 2, we need to sum the probabilities that x is 1 or 2.

$$P(X \leq 2) = P(x = 1) + P(x = 2) = \frac{1}{14} + 4(\frac{1}{14}) = \frac{5}{14}$$

Now we will find E[X] by summing the product of each outcome by it's probability.

$$E[X] = 1 \times \frac{1}{14} + 2 \times \frac{4}{14} + 3 \times \frac{9}{14} = \frac{36}{14} = 2.57$$

With E[X], we can calculate Var[X]:

$$Var[X] = (1 - 2.57)^2 \times \frac{1}{14} + (2 - 2.57)^2 \times \frac{4}{14} + (3 - 2.57)^2 \times \frac{9}{14} = 0.3877571$$

2. Suppose that X is a continuous random variable having p.d.f. $f(x) = cx^2$ for 1 ≤ x ≤ 2. Find c, Pr(X ≥ 1), E[X], and Var(X).

$$\int_1^2 cx^2dx = \left[\frac{cx^3}{3}\right]_1^2$$
$$\frac{8c}{3} - \frac{c}{3} = 1$$
$$c = \frac{3}{7}$$

Since the range of value is 1 ≤ x ≤ 2, the P(x ≥ 1) = 1.

$$E[X] = \int_a^b xf(x)dx$$

$$E[X] = \int_1^2 x\frac{3x^2}{7}dx = \left[\frac{3x^4}{28} + C\right]_1^2 = (\frac{48}{28} + C) - (\frac{3}{28} + C) = \frac{45}{28} = 1.607$$

$$Var[X] = E[X^2] - E[X]^2 = \int_1^2 x^2\frac{3x^2}{7}dx - (\frac{45}{28})^2 = \left[\frac{3x^5}{35} + C\right]_1^2 - (\frac{45}{28})^2 = \frac{96}{35} - \frac{3}{35} - \frac{2025}{784} = 0.0742$$

3. Suppose that X and Y are jointly continuous random variables with: $y - x$ for $0 < x < 1$ and $1 < y < 2$, otherwise 0

a. Compute and plot $f_X(x)$ and $f_Y(y)$.

$$for\ 0 < x < 1,\ \ f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

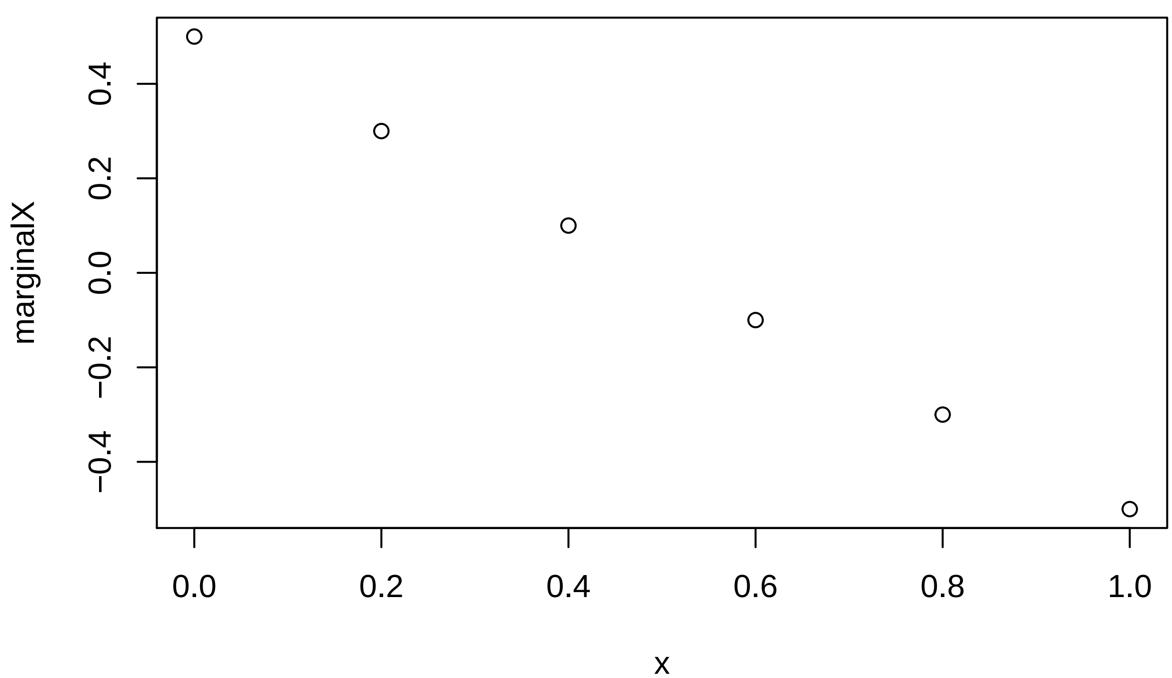
$$f_X(x) = \int_0^1 y - xdy = \left[\frac{y^2}{2} - xy\right]_0^1$$

$$f_X(x) = \frac{1}{2} - x$$

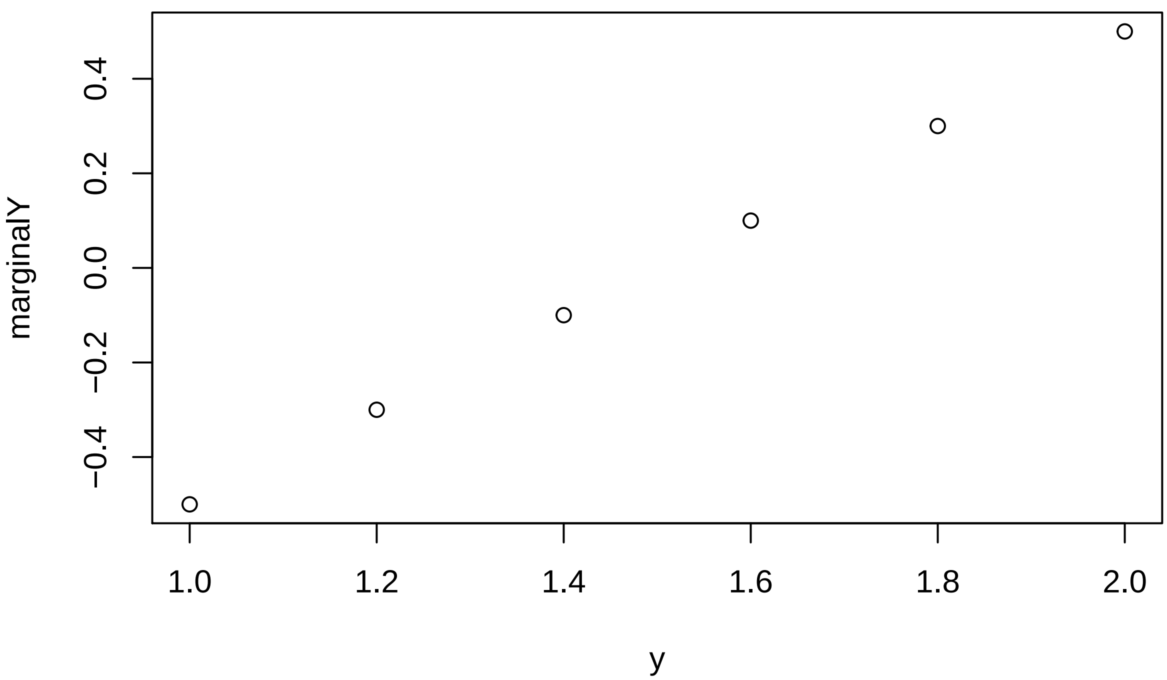
$$for\ 1 < y < 2,\ \ f_Y(y) = \int_1^2 y - x\,dx = \left[xy - \frac{x^2}{2}\right]_1^2$$

$$f_Y(y) = (2y - 2) - (y - \frac{1}{2}) = y - \frac{3}{2}$$

```
x <- seq(0,1,by=0.2)
fxx <- function(x){0.5-x}
marginalX <- fxx(x)
y <- seq(1,2,by=0.2)
fyy <- function(y){y-1.5}
marginalY <- fyy(y)
plot(x,marginalX)
```



```r
plot(y,marginalY)
```



b. Are X and Y independent?

If $E[XY] = E[X]E[Y]$, then X and Y are independent.

From the solutions below, we know that $E[X] = \frac{-1}{12}$, $E[Y] = \frac{1}{12}$, and $E[XY] = \frac{2}{3}$. Since $\frac{2}{3} \neq \frac{-1}{144}$, we know that X and Y are not independent.

c. Compute $F_X(x)$ and $F_Y(y)$.

$$f_X(x) = \frac{dF_X x}{dx}$$

$$f_X(x) = \frac{1}{2} - x$$

$$F_X x = \int_{-\infty}^{\infty} \frac{1}{2} - x = \frac{x - x^2}{2} + C$$

$$f_Y(y) = y - \frac{3}{2}$$

$$F_Y(y) = \int_{-\infty}^{\infty} y - \frac{3}{2} = \frac{y^2 - 3y}{2} + C$$

d. Compute E[X], Var(X), E[Y], Var(Y), Cov(X,Y), and Corr(X,Y).

$$E[X] = \int_0^1 u f_x(u) du = \int_0^1 u(\frac{1}{2} - u) du = \int_0^1 \frac{u}{2} - u^2 = \left[\frac{u^2}{4} - \frac{u^3}{3}\right]_0^1 = \frac{1}{4} - \frac{1}{3} = -\frac{1}{12}$$

$$E[X] = -\frac{1}{12}$$

$$Var[X] = \int_0^1 u^2 f_x(u)du - (\frac{-1}{12})^2$$

$$Var[X] = \int_0^1 \frac{u^2}{2} - u^3 - \frac{1}{144}$$

$$Var[X] = \left[\frac{u^3}{6} - \frac{u^4}{4}\right]_0^1 - \frac{1}{144}$$

$$Var[X] = \frac{1}{6} - \frac{1}{4} - \frac{1}{144} = \frac{24}{144} - \frac{36}{144} - \frac{1}{144} = \frac{-13}{144}$$

$$E[Y] = \int_1^2 uf_y(u)du = \int_1^2 u(u - \frac{3}{2})du = \int_1^2 u^2 - \frac{3u}{2}du$$

$$E[Y] = \left[\frac{u^3}{3} - \frac{3u^2}{4}\right]_1^2$$

$$E[Y] = (\frac{8}{3} - \frac{12}{4}) - (\frac{1}{3} - \frac{3}{4})$$

$$E[Y] = \frac{1}{12}$$

$$Var[Y] = \int_1^2 u^2 f_y(u)du - (\frac{1}{12})^2$$

$$Var[Y] = \int_1^2 u^2(u - \frac{3}{2})du - \frac{1}{144}$$

$$Var[Y] = \int_1^2 u^3 - \frac{3u^2}{2}du - \frac{1}{144}$$

$$Var[Y] = \left[\frac{u^4}{4} - \frac{u^3}{2}\right]_1^2 - \frac{1}{144}$$

$$Var[Y] = (\frac{16}{4} - \frac{8}{2}) - (\frac{1}{4} - \frac{1}{2}) - \frac{1}{144}$$

$$Var[Y] = \frac{35}{144}$$

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

$$E[XY] = \int_0^1 \int_1^2 xy(y - x)dydx = \int_0^1 \int_1^2 xy^2 - x^2y\, dydx$$

$$E[XY] = \int_1^2 \left[\frac{x^2y^2}{2} - \frac{x^3y}{3}\right]_0^1 dy$$

$$E[XY] = \int_1^2 \frac{y^2}{2} - \frac{y}{3}dy = \left[\frac{y^3}{6} - \frac{y^2}{6}\right]_1^2$$

$$E[XY] = (\frac{8}{6} - \frac{4}{6}) - (\frac{1}{6} - \frac{1}{6}) = \frac{2}{3}$$

$$Cov(X, Y) = \frac{2}{3} - \frac{-1}{12} \times \frac{1}{12}$$

$$Cov(X,Y) = \frac{49}{144} = 0.34$$

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\frac{49}{144}}{\sqrt{\frac{-13}{144} \times \frac{35}{144}}}$$

4. Suppose that the following 10 observations come from some distribution (not highly skewed)with unknown mean μ.

```r
sample <- c(7.3,6.1,3.8,8.4,6.9,7.1,5.3,8.2,4.9,5.8)
mean(sample)
```

```
## [1] 6.38
```

```r
var(sample)
```

```
## [1] 2.161778
```

The sample mean is $\bar{X} = 6.38$ and the variance is $S^2 = 2.161778$.

We know that 95% of the values in a normally distributed population will fall within 2 standard deviations of the mean. We can construct the 95% confidence interval for μ as follows:

$$\bar{x} \pm 2\sigma_x$$
$$6.38 \pm 2\sqrt(\frac{2.161778}{10})$$
$$6.38 \pm 0.93 = (5.45, 7.31)$$

5. A random variable X has the memoryless property if, for all $s, t > 0$, $P(x > t + s \mid x > t) = P(x > s)$.

Show that the exponential distribution has the memoryless property.

We can rearrange the above equivalency as follows:

$$\frac{P(x > t + s, x > t)}{P(x > t)} = P(x > s)$$
$$P(x > s + t) = P(x > s)P(x > t)$$

If we take the exponential distribution as: $f(x, \lambda) = \lambda e^{-\lambda x}$, then we can see that the equation holds true for all values of s as long as t $\neq 0$. This equality makes sense when we consider a scenario with waiting times where x is the time it takes for an event to occur. If an event fails to occur within the initial time s, then the probability of the event occuring in the initial time t should be equal to the unconditioned probability (the right-hand side).

$$P(x > t) = e^{-\lambda t}$$
$$P(x > s) = e^{-\lambda s}$$
$$P(x > s + t) = e^{-\lambda s}e^{-\lambda t} = e^{-\lambda(s+t)}$$

6. Suppose X1,X2,...,Xn are i.i.d. $Exp(\lambda = 1)$. Use the Central Limit Theorem to find the approximate value of $P(100 \leq \sum_{i=1}^{100} X_i \leq 110)$.

First, we know that the mean and the variance of an exponentially distributed random with $\lambda = 1$ are:

$$E[X] = \frac{1}{\lambda} = 1$$

$$Var[X] = \frac{1}{\lambda^2} = 1$$

From the Central Limit Theorem, we know that the sample mean, $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$. Since $n = 100$, we can rewrite $P(100 \leq \sum_{i=1}^{100} X_i \leq 110)$ as:

$$P(1 \leq \bar{X} \leq 1.1) = F(1.1) - F(1)$$

We determine the probability by finding the z-scores that correspond to the standard normal distribution.

$$z = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{1.1 - 1}{\frac{1}{\sqrt{100}}} = 1$$

The area under the curve to the left of z=1 is 0.8413.

$$F(1.1)z = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{1 - 1}{\frac{1}{\sqrt{100}}} = 0$$

The area under the curve to the left of z=0 is 0.5000.

$$P(1 \leq \bar{X} \leq 1.1) = 0.8413 - 0.5000 = 0.3413$$

Exercises from Discrete-Event System Simulation:

5.13

A random variavle X that has a pmf given by $p(x) = 1/(n + 1)$ over the range $R_X = 0, 1, 2, ..., n$ is said to have a discrete uniform distribution.

a. Find the mean and variance of this distribution.

The mean of the distribution is:

$$E[X] = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} x_i = \frac{n(n+1)}{2}$$

$$E[X] = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$$

$$Var[X] = E[X^2] - E[X]^2$$

$$E[X^2] = \frac{1}{n}\sum_{i=1}^{n} x_{i^2}$$

$$\sum_{i=1}^{n} x_{i^2} = \frac{n(n+1)(2n+1)}{6}$$

$$E[X^2] = \frac{1}{n} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$Var[X] = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$Var[X] = \frac{n^2-1}{12}$$

b. If $R_X = a, a+1, a+2, ...b$, compute the mean and variance of X.

If the discrete uniform distribution has parameters a and b where $a < b$, then the probability mass function is $f(x) = \frac{1}{b-a+1}$. The solution for E[X] above shows use that we should expect the mean to fall in the middle of a discrete uniform distribution. If our interval is (a,b), we can intuit that the mean should be:

$$E[X] = \frac{a+b}{2}$$

We also know that the variance will behave the same over the interval (a, b), so we can plug in $n = b - a + 1$ in $Var[X] = \frac{n^2-1}{12}$:

$$Var[X] = \frac{(b-a+1)^2 - 1}{12}$$

5.14

The lifetime, in years, of a satellite placed in orbit is given by the following pdf:

$$f(x) = 0.4e^{-0.4x} \; when \; x \geq 0, \; otherwise \; 0$$

$$P(x > t) = \int_t^{\infty} \lambda e^{-\lambda x} dx = \left[\lambda \frac{-e^{-\lambda x}}{\lambda}\right]_t^{\infty} = e^{-\lambda t}$$

a. What is the probability that this satellite is still "alive" after 5 years?

$$P(x > 5) = e^{-(0.4)(5)} = 0.135$$

b. What is the probability that the satellite dies between 3 and 6 years from the time it is placed in orbit?

$$P(3 \leq x \leq 6) = F(6) - F(3) = (1 - e^{-(0.4)(6)}) - (1 - e^{-(0.4)(3)})$$
$$P(3 \leq x \leq 6) = 0.210$$

5.39

Three shafts are made and assembled into a linkage. The length of each shaft, in centimeters, is distributed as follows:

Shaft 1: $N(60, 0.09)$ Shaft 2: $N(40, 0.05)$ Shaft 3: $N(50, 0.11)$

a. What is the distribution of the length of the linkage?

The sum of multiple mutually independent normal random variables also has a normal distribution. The mean and variance of the distribution can be summed from the individual distributions.

$$E[X] = 60 + 40 + 50 = 150$$
$$Var[X] = 0.09 + 0.05 + 0.11 = 0.25$$

The distribution of the linkage is $N(150, 0.25)$.

b. What is the probability that the linkage will be longer than 150.2 centimeters?

Here we will use a z-table to find $P(X > 150.2)$.

$$z = \frac{150.2 - 150}{\sqrt{0.25}} = 0.4$$
$$P(X > 150.2) = 1 - P(X \leq 150.2)$$
$$= 1 - 0.6554$$
$$= 0.3446$$

The probability that the linkage is longer than 150.2 cm is 0.3446.

c. The tolerance limits for the assembly are (149.83,150.21). What proportion of assemblies are within the tolerance limits?

We will use z-scores again the portion of the assemblies that fall within the tolerance limits.

$$P(149.83 \leq X \leq 150.21) = P(X \leq 150.2)) - P(X \leq 149.83)$$
$$z = \frac{150.21 - 150}{\sqrt{0.25}} = 0.42$$

The area under the curve to the left of $z = 0.42$ is 0.6628.

$$z = \frac{149.83 - 150}{\sqrt{0.25}} = -0.34$$

The area under the curve to the left of $z = -0.34$ is 0.3669.

$$0.6628 - 0.3669 = 0.2959$$

The proportion of assemblies that fall within the tolerance limits is 0.2959.