CS 5135/6035 Learning Probabilistic Models

Exercise Solutions for Lecture 4 (Discrete Probability Distributions)

Gowtham Atluri

9/6/2018

Questions

1. Prove that $E(aX+b) = aE(X)+b$, where X is a random variable, a and b are constants. [**2 points**]
   Solution:

$$\begin{aligned}
E(aX+b) &= \sum_{X=1}^{n} (aX+b)p(X) \\
&= \sum_{X=1}^{n} aXp(X) + bp(X) \\
&= \sum_{X=1}^{n} aXp(X) + \sum_{X=1}^{n} bp(X) \\
&= a\sum_{X=1}^{n} Xp(X) + b\sum_{X=1}^{n} p(X) \\
&= aE(X) + b, \text{ since} \sum_{X=1}^{n} p(X) = 1
\end{aligned}$$

2. Write the cumulative distribution function for the probability distribution [**3 points**]

$$\begin{aligned}
p(x=1) &= 0.1 \\
p(x=2) &= 0.3 \\
p(x=3) &= 0.4 \\
p(x=4) &= 0.2
\end{aligned}$$

   Solution:

$$cdf(x) = \begin{cases} 0, & x < 1 \\ 0.1, & 1 \le x < 2 \\ 0.4, & 2 \le x < 3 \\ 0.8, & 3 \le x < 4 \\ 1, & x \ge 4 \end{cases}$$

3. In a study of battery life for laptop computers, researchers found that the probability that the battery life (L) will exceed 5 hours is 0.12. Three such batteries are used in independent laptops and we are interested in finding the probability that some x of the three batteries will last 5 hours or more. [**10 points**]

   (i) Specify the standard probability distribution you will use to model this scenario.
   (ii) What is the state space or the set of possible outcomes for this scenario.
   (iii) Find the probability that only one of the three batteries will last 5 hours or more.
   (iv) Write Julia code to plot the *true* distribution for all possible values of x.

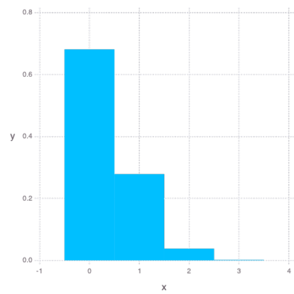

Figure 1: Result for Q3 (iv)

(v) Write Julia code to sample *1000* points from this distribution and plot the *empirical* distribution. [*Hint:* True distribution will contain the probabilities from the probability function (using *pdf*() in Julia), whereas the empirical distribution will have probabilities computed from the samples of the distribution.]

Solution:

(1) Binomial Distrbution

(ii) ${ 0,1,2,3}

(iii)

$$p(x=1) = \binom{3}{1}(0.12)^1(1-0.12)^{3-1}$$
$$= 0.2787$$

(iv)

```
using Distributions, Gadfly;
d = Binomial(3,0.12);
x = collect(0:3);
y = pdf.(d,x);
myplot = plot(x=x, y=y, Geom.bar);
draw(PNG("./q3vi.png", 5inch, 5inch), myplot);
```

(v)

```
using Distributions, Gadfly;
d = Binomial(3,0.12);
data = rand(d,1000);
myplot = plot(x=data, Geom.histogram(bincount=4));
draw(PNG("./q3v.png", 5inch, 5inch), myplot);
```

4. A statistically inclined farmer would like to model the number of grasshoppers per square meter of his rangeland. He is told that typically there are 0.5 grasshoppers per square meter on a rangeland. **[10 points]**

(i) Specify the standard probability distribution you will use to model this scenario.
(ii) What is the state space or the set of possible outcomes for this scenario.
(iii) Find the probability that there are five or more grasshoppers in a randomly selected square meter region.
(iv) Write Julia code to plot the *true* distribution for values of x upto 10.
(v) Write Julia code to sample *1000* points from this distribution and plot the *empirical* distribution.
Solution:

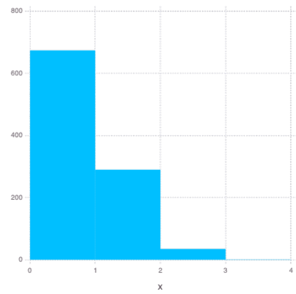

Figure 2: Result for Q3 (v)

(1) Poisson Distrbution

(ii) \${ 0,1,2,3, . . . }

(iii)

$$
\begin{aligned}
p(x) &= \frac{\lambda^x}{x!}e^{-\lambda} && \lambda = 0.5 \\
p(x=0) &= \frac{0.5^0}{0!}e^{-0.5} && = 0.6065 \\
p(x=1) &= \frac{0.5^1}{1!}e^{-0.5} && = 0.3033 \\
p(x=2) &= \frac{0.5^2}{2!}e^{-0.5} && = 0.0758 \\
p(x=3) &= \frac{0.5^3}{3!}e^{-0.5} && = 0.0126 \\
p(x=4) &= \frac{0.5^4}{4!}e^{-0.5} && = 0.0016 \\
p(x \geq 5) &= 1 - \sum_{x=0}^{4} p(x) && = 1 - 0.9998 = 0.0002
\end{aligned}
$$

(iv)

```
using Distributions, Gadfly;
d = Poisson(0.5);
x = collect(0:10);
y = pdf.(d,x);
myplot = plot(x=x, y=y, Geom.bar);
draw(PNG("./q4vi.png", 5inch, 5inch), myplot);
```

(v)

```
using Distributions, Gadfly;
d = Poisson(0.5);
data = rand(d,1000);
myplot = plot(x=data, Geom.histogram(bincount=10));
draw(PNG("./q4v.png", 5inch, 5inch), myplot);
```

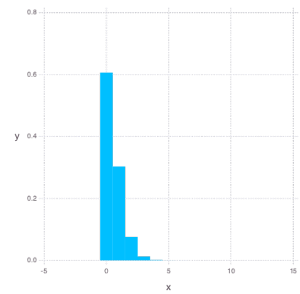

Figure 3: Result for Q4 (iv)

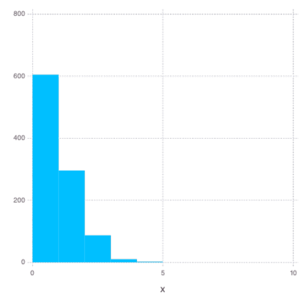

Figure 4: Result for Q4 (v)

Bonus Questions

1. **Matching coins** A and B agree to the following rules of the coin-matching game: A wins \$1 if the match is tails and \$2 if the match is heads. A loses \$1 if coins do not match. Write the probability distribution of A's winning amount per game. What is A's expected winning ammount? Is this a fair game?
2. Prove that $V(aX+b) = a^2V(X)$, where X is a random variable and a is a constant.
3. Prove that $p(x \geq j+k|X \geq j) = p(x \geq k)$ for a geometric distribution.
4. A large lot of tires contains 5% defectives. Four tires are to be chosen from the lot and placed on a car.
   (i) Find the probability that two defectives are found before four good ones.
   (ii) Find the expected value and variance of the number of selections that must be made to get four good tires. (*Hint:* First find the expected value and variance of the number of defective tires that will be selected before finding the four good ones.)