ECE4010J RC Week 3

Linda Wei

I Things need to be mentioned

1. ECE4010J, engineering course. Application rather than Math is important, the RC I conduct in this course is similar to CHEM2100J instead of MATH2030J or MATH2560J. (MMA, not focus on pure math)
2. HMWK, online module and quiz. The homework will not be counted in your final grade, however, online module and quiz do. So, please pay attention to try your best to get higher score in OM and quiz.
3. Piazza, asking questions is very important and piazza is a good place. Though I'm very busy this semester, there are other TAs and students who can answer your questions.
4. Catme. This is a website where you do your peer evaluation, do not forget it, it will cost a lot of points. When there is a peer review, an email will be sent to your mail-box, pay attention to this mail, pay attention to your trash bin
5. Project HC. This course lays much stress on HC violation, pay attention to make sure that you and your teammates do not copy others' work! I can not save you if you violate HC like what some of you attempted to do in MATH2030J and MATH2560J.
6. MMA,MMA and MMA! In my RC, you are information processors who feed

useful information to MMA, you are not the problem solver, MMA is!

II. Distributions

1. Bernoulli Distribution

Example: Tossing a coin follows a Bernoulli Distribution. In this distribution, we denote the probability of "success" (x=1) as p, "failure" as 1-p.

PDF: $f_X(x) = p^x(1-p)^{1-x}$. This form is just a fancy form of writing
Mathematica: Bernoulli Distribution[p].
Expectation: p
Variance:(1-p)p

2. Binomial Distribution

Example: Tossing a coin for many times. Binomial Distribution is the sum of some independent Bernoulli Distribution. It represents the experiments with total number n and successful probability p.

PDF: $f_X(x) = \mathrm{Binomial}(n,x)p^x(1-p)^{n-x}$
Mathematica: Bernoulli Distribution[n,p]
Expectation: np
Variance:n(1-p)p

Warm up exercise: Tossing an unfair coin, which has a 0.3 possibility of turning up. What is the probability of show up at least 18 times if you toss 50 times?
Solution: Do not use your hand if necessary in this course!!!! If you want to plainly get higher grade.

```
In[10]:= Clear[x]
         清除

In[14]:= Probability[x ≥ 18, x ≈ BinomialDistribution[50, 0.3]]
         概率                    二项分布

Out[14]= 0.217807
```

One line of code, simple! The “follows” is esc+dist

Past year exam problem exercise:
1.
A coin is tossed 28800 times. Assuming that the coin is fair (i.e., heads and tails each have a 50% chance of coming up), give the probability that between 14380 and 14399 heads come up. (we will talk about using the normal approximation to the binomial distribution later, but purely use binomial distribution is a way to verify your answer.)

Information processor: n=28800, p=0.5, 14380<=x<=14399
Problem Solver:

```
Clear[x]
清除

Probability[14 380 ≤ x ≤ 14 399,
概率
 x ≈ BinomialDistribution[28 800, 0.5]]
     二项分布

0.0931019
```

2.
Suppose that a hotel has 200 equally-sized rooms. From experience, it is known that 10% of travellers who have reserved a room do not turn up. Answer the following questions.

i) Suppose that the hotel has received reservations for 215 rooms. What is the probability that there will enough rooms for all travelers who turn up?

ii) ii) Suppose that the hotel has received reservations for 215 rooms. What is the probability that more than 190 rooms will be occupied? (This question seems that it do not care whether there is enough rooms)

iii) iii) At what point should the hotel stop accepting reservations if management wants to be at least 99% certain to have enough rooms for all travelers who show up?

I) Information processor: p=1-0.1=0.9 , n=215, x<=200.

Problem solver:

```
In[10]:= Clear[x]
         清除

In[16]:= Probability[x ≤ 200, x ≈ BinomialDistribution[215, 0.9]]
         概率                    二项分布

Out[16]= 0.950358
```

II) Information processor: p=1-0.1=0.9 , n=215, 190<x<=200.

Problem solver:

```
In[10]:= Clear[x]
         清除
In[18]:= Probability[190 < x, x ≈ BinomialDistribution[215, 0.9]]
         概率                     二项分布
Out[18]= 0.757507
```

III) Information processor: p=1-0.1=0.9 , n=???, x<=200.

Probability[x<=200,x follows Binomial[n,0.9]] is >=0.99, by trial

```
In[26]:= Probability[x ≤ 200, x ≈ BinomialDistribution[212, 0.9]]
         概率                     二项分布
Out[26]= 0.991277
In[27]:= Probability[x ≤ 200, x ≈ BinomialDistribution[213, 0.9]]
         概率                     二项分布
Out[27]= 0.983321
```

It seems 212.

Seems easy? Right?

Exercise 9.
Let X be a discrete random variable following a Bernoulli distribution with parameter $p = 1/2$ and let $X_1, \dots, X_{10}$ be a random sample of size 10. Calculate the probability that the sample mean is greater than 3/4, i.e., find

$$P[\overline{X} > 3/4].$$

We note that $X_1 + \cdots + X_{10}$ follows a binomial distribution with $n = 10$ and $p = 1/2$.

This observation is worth 2 Marks.

Then

$$\begin{aligned} P[\overline{X} > 3/4] &= P[X_1 + \cdots + X_{10} > 3/4 \cdot 10] \\ &= P[X_1 + \cdots + X_{10} > 7.5] \\ &= 1 - P[X_1 + \cdots + X_{10} < 7.5] \\ &= 1 - P[X_1 + \cdots + X_{10} \le 7] \\ &= 0.0547 \end{aligned}$$

Give a total of 2 Marks for the calculation, subtract 1 Mark if the final result is incorrect. Give 0 marks if there is no calculation.

```
In[32]:= Probability[x > 0.75 * 10, x ≈ BinomialDistribution[10, 0.5]]
         概率                         二项分布
Out[32]= 0.0546875
```

3. Geometric Distribution

Meaning: the times that first success happens.

Attention: the definition on slide is not the definition on the MMA, the MMA defines it as the number of failure that is needed!!!

PDF: $f_X(x) = p(1-p)^{x-1}$
Mathematica: GeometricDistribution[p],pay attention when you use.
Expectation: 1/p
Variance:(1-p)/p^2

4. Pascal Distribution

Meaning: the number of trial x before the nth success.

It is the sum of independent geometric distribution

PDF: $f_X(x) = \text{Binomial}(x-1,n-1)p^x(1-p)^{(n-x)}$

Mathematica: PascalDistribution[n,p]

Past time exam exercises:

A fair coin is tossed repeatedly.
i) What is the probability that the 5th head occurs on the 10th toss?
ii) Suppose that we know that the 10th head occurs on the 25th toss. Find the probability density function of the toss number of the 5th head

i)

Information processor: x=10, n=5 , p=0.5

Problem Solver:

```
In[29]:= Clear[x]
         清除
In[31]:= Probability[x == 10, x ≈ PascalDistribution[5, 0.5]]
         概率                     帕斯卡分布
Out[31]= 0.123047
```

ii) The tenth head occurs on the 25^{th} toss, so nine heads occur in the first 24 tosses. There are $\binom{24}{9}$ ways of selecting the toss numbers for these nine heads. If the 5^{th} head occurs on the x^{th} toss, then 4 heads must have occurred in the first $x-1$ tosses and 4 heads must occur in the following $24-x$ tosses. Here $5 \le x \le 20$. It follows that $f_X : \Omega \to \mathbb{R}$ given by

$$f_X(x) = P[X = x] = P[5^{\text{th}} \text{ head on } x^{\text{th}} \text{ toss} \mid 10^{\text{th}} \text{ head on } 25^{\text{th}} \text{ toss}] = \frac{\binom{x-1}{4}\binom{24-x}{4}}{\binom{24}{9}}$$

is the density of the random variable $X : S \to \Omega = \{5, 6, \dots, 20\}$, where X is the toss number of the 5^{th} head and S is the sample space for the experiment.
(3 Marks)

Exercise2

Let X be a discrete random variable following a geometric distribution with parameter $p = 1/2$ and let $X_1, \dots, X_{10}$ be a random sample of size 10. Calculate the probability that the sample mean is no more than than 1.5, i.e., find

$$P[\overline{X} \le 1.5].$$

(4 Marks)

We note that $X_1 + \cdots + X_{10}$ follows a Pascal distribution with $r = 10$ and $p = 1/2$.

This observation is worth 2 Marks.

Then

$$\begin{aligned}
P[\overline{X} < 1.5] &= P[X_1 + \cdots + X_{10} \le 10 \cdot 1.5] \\
&= P[X_1 + \cdots + X_{10} \le 15] \\
&= \sum_{x=10}^{15} \binom{x-1}{9} \frac{1}{2^x} \\
&= \frac{1}{1024}\left(\binom{9}{9} + \frac{1}{2}\binom{10}{9} + \frac{1}{4}\binom{11}{9} + \frac{1}{8}\binom{12}{9} + \frac{1}{16}\binom{13}{9} + \frac{1}{32}\binom{14}{9} + \frac{1}{64}\binom{15}{9}\right) \\
&= \frac{1}{1024}\left(1 + \frac{10}{2} + \frac{55}{4} + \frac{220}{8} + \frac{715}{16} + \frac{2002}{32} + \frac{5005}{64}\right) \\
&= \frac{309}{2048} = 0.15
\end{aligned}$$

Give a total of 2 Marks for the calculation, subtract 1 Mark if the final result is incorrect. Give 0 marks if there is no calculation.

```
In[34]:= Probability[x ≤ 15, x ≈ PascalDistribution[10, 0.5]]
         概率                      帕斯卡分布

Out[34]= 0.150879
```

5. Poisson Distribution

Poisson Distribution describes the occurrence of events that occur at a constant rate in a continuous environment.

PDF $f_X(x)$= $e^{-\mu} \mu^x / x!$

Expectation: u

Variance:u

6. Exponential Distribution

PDF $f_X(x)= be^{-bt}$

Expectation : 1/b

Variance:1/b^2

Exercise 5.
A certain widget has a mean time between failures of 24 hours, i.e., failures occur at a constant rate of one failure every 24 hours.

One evening, the widget was observed to be working at 10 pm and then left unobserved for the night. The next morning at 6 am, it was observed to have failed earlier. What is the probability that it was still working at 5 am that morning?
(5 Marks)

Solution 5.
The failures of the widget follow a Poisson distribution with a rate $\lambda = 1/24$. The time to failure is therefore exponentially distributed with parameter $\beta = \lambda = 1/24$.
Generally, the failure density is $f_\beta(t) = \beta e^{-\beta t}$ so that

$$P[\text{widget fails between time } T_1 \text{ and } T_2] = \int_{T_1}^{T_2} \beta e^{-\beta t}\, dt$$

Therefore,

$$\begin{aligned}
&P[\text{widget fails between time } T_1 \text{ and } T_2 \mid \text{widget has failed not after time } T_2] \\
&= \frac{P[\text{widget fails between time } T_1 \text{ and } T_2]}{P[\text{widget has failed not after time } T_2]} \\
&= \frac{\int_{T_1}^{T_2} \beta e^{-\beta t}\, dt}{\int_0^{T_2} \beta e^{-\beta t}\, dt} \\
&= \frac{-e^{-\beta t}|_{T_1}^{T_2}}{-e^{-\beta t}|_0^{T_2}} \\
&= \frac{e^{-\beta T_1} - e^{-\beta T_2}}{1 - e^{-\beta T_2}}
\end{aligned}$$

Inserting our values, the probability is

$$\frac{e^{-7/24} - e^{-1/3}}{1 - e^{-1/3}} = 0.108.$$

i) *2 Marks for writing down the correct probability that the widget fails between time T_1 and T_2.*

ii) *2 Marks for the calculation using conditional probabiliy/Bayes's theorem.*

iii) *1 Mark for the correct correct result, if it is supported by calculation above.*

7. Normal Distribution

1. Use Normal approximate to solve the 1st problem in Binomial distribution

Solution 6.
We want to approximate

$$\mathcal{P} = \sum_{k=14380}^{14399} \binom{28800}{k} 0.5^k 0.5^{28800-k}.$$

We use the normal approximation, so that

$$\mu = np = 28800 \cdot 0.5 = 14400$$

and

$$\sigma^2 = np(1-p) = 0.5 \cdot 28800 \cdot (1-0.5) = 7200.$$

(1 Mark) Note that $\sigma = 60\sqrt{2}$. Taking into account the half-unit correction, **(1 Mark)** we have ($Z = (X-14400)/(60\sqrt{2})$)

$$\begin{aligned}
\mathcal{P} &\approx P[14379.5 \le X \le 14399.5] \\
&= P\left[\frac{14379.5-14400}{60\sqrt{2}} \le Z \le \frac{14399.5-14400}{60\sqrt{2}}\right] \\
&= P[-0.24 \le Z \le -0.0059] \\
&= P[0 \le Z \le 0.24] - P[0 \le Z \le 0.0059] \\
&= 0.0948 - 0.0000 = 0.0948 = 9.48\%.
\end{aligned}$$

Exercise 11.
The target value for the thickness of a machined cylinder is 8 cm. The upper specification limit is 8.2 cm and the lower specification limit is 7.9 cm. A machine produces cylinders that have a mean thickness of 8.1 cm and a standard deviation of 0.1 cm. Assume that the thickness of the cylinders produced by this machine follows a normal distribution.

i) What is the probability that the thickness of a randomly selected cylinder is within specification?

ii) What is the probability that the thickness of a randomly selected cylinder exceeds the target value?

iii) Find a bound $L > 0$ such that the thickness of 90% of all cylinders lies within $(8.1 \pm L)$ cm.

(2+2+2 Marks)

2.

Solution 11.

i) Denote the random variable "thickness" by T. Then the variable $Z = \frac{T-8.1}{0.1}$ is standard normal and

$$\begin{aligned}
P[7.9 \le T \le 8.2] &= P\left[\frac{7.9-8.1}{0.1} \le \frac{T-8.1}{0.1} \le \frac{8.2-8.1}{0.1}\right] \\
&= P\left[\frac{7.9-8.1}{0.1} \le Z \le \frac{8.2-8.1}{0.1}\right] \\
&= P[-2 \le Z \le 1] \\
&= P[0 \le Z \le 1] + P[0 \le Z \le 2] \\
&= 0.3413 + 0.4772 = 0.8185 \approx 82\%.
\end{aligned}$$

(2 Marks)

ii) We have

$$\begin{aligned}
P[T \ge 8] &= P\left[\frac{T-8.1}{0.1} \ge \frac{8-8.1}{0.1}\right] \\
&= P[Z \ge -1] \\
&= P[0 \le Z \le 1] + P[0 \le Z] \\
&= 0.3413 + 0.5 = 0.8413 \approx 84\%.
\end{aligned}$$

(2 Marks)

iii) We want to find L such that

$$0.05 = P[T > 8.1 + L] = P\left[\frac{T-8.1}{0.1} > 10L\right] = 1 - P[Z \le 10L] = 0.5 - P[0 \le Z \le 10L]$$

(1 Mark) From the table, $10L = 1.645$ (interpolating between 1.64 and 1.65), so $L = 0.1645$. Thus the thickness of 90% of all cylinders lies within (8.1 ± 0.16) cm. **(1 Mark)**

Complimentary exercise (wont be tested, search online and solve them)

1.

Given the PDF of two distribution p(x),q(x). We define the quantity:

$$D_{KL}(p||q) = \mathbb{E}_{p(x)}[\log(p(x)/q(x))] = \int p(x)\log(\frac{p(x)}{q(x)})dx$$

as the Kullback–Leibler(KL)-Divergence (or relative entropy) of the distribution p and q. Consider the following problems:

1. Show that: If p and q are normal distributions that are of mean and standard deviation $[\mu_1, \sigma_1]$, $[\mu_2, \sigma_2]$. $D_{KL}(p||q) = \log(\frac{\sigma_2}{\sigma_1}) + \frac{\sigma^2+(\mu_1-\mu_2)^2}{2\sigma_2^2} - \frac{1}{2}$
2. Show that: given arbitrary non-zero distribution p and q, $D_{KL}(p||q) \geq 0$

2.

Given the conditional probability density function $q(x_t \mid x_{t-1}) = N(\sqrt{1-\beta_t}x_{t-1}, \beta_t)$, $t \in \mathbb{N}$. β_t is a constant value in (0,1) if t is fixed Assume in a certain engineering problem, we want to sample x_t from $q(x_t \mid x_{t-1})$, here we do not directly sample from that distribution, instead, we play the following trick, steps follows:

1. We sample z_{t-1} from the Standard Normal Distribution $[0, 1]$
2. Then we calculate $x_t = z_{t-1} * \beta_t + \sqrt{1-\beta_t}x_{t-1}$
3. You can see that we have sampled the x_t that follows the requirment.

Now, let $\alpha_t = 1 - \beta_t$consider the following problems:

1. Show that $x_t = \sqrt{a_t a_{t-1}}x_{t-2} + \sqrt{1 - a_t a_{t-1}}\bar{z_{t-2}}$ in which $\bar{z_{t-2}}$ is another sample from the standard normal distribution.
2. Show that $x_t = \sqrt{\bar{a}_t}x_0 + \sqrt{1-\bar{a}_t}z$, in which z is another sample from the standard normal distribution.
3. Consider what will happen to $q(x_t \mid x_0)$ if β_t is increasing when t becomes large.

Remark: The above sample technique is called reparameterization in real engineering problem. You can view VAE model and Diffusion model in deep learning if you want to know its merit.