Math-233: Theory of probability: Oct 28th Notes

Kaichen Ma, Graham Harris

October 31st 2020

1 Introduction

This paper covers all materials from both the lecture video and class recording from Prof. Mike O'Neil's October 28th lecture for Theory of Probability, with some additional sources from the textbook.

2 Asynchronous Lecture

2.1 Uniform Random Variable

A uniform random variable is one that takes on values in an interval with equal probability.
For instance, a random variable is said to be uniformly distributed over the interval (0, 1) if its probability density function is given by

$$U \sim uniform(0,1)$$

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Assume real numbers a,b in interval (0,1) and a<b, then

$$\begin{aligned} P\left[U \in (a,b)\right] &= \int_a^b f(x)dx & (2)\\ &= \int_a^b dx & (3)\\ &= b - a & (4) \end{aligned}$$

Furthermore, the cumulative distribution function can be written as

$$F(x) = P[U \le x] \tag{5}$$

$$= \begin{cases} 0 & x \le 0 \\ x & x \in (0,1) \\ 1 & x \ge 1 \end{cases} \tag{6}$$

In general, the uniform density can be expanded to define on any interval (α, β)

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{for } u \in (\alpha, \beta) \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Similarly, to compute the probability of sub-interval (a, b) where $\alpha < a < b < \beta$ can be formulated as

$$P[U \in (a,b)] = \int_a^b \frac{1}{\beta - \alpha} dx \tag{8}$$

$$= \frac{b-a}{\beta-\alpha} \tag{9}$$

(the outcome of this integral is not exceeding 1)
Since $F(a) = \int_{-\infty}^{a} f(x)dx$ that the distribution function of a uniform random variable on the interval (α, β) is given by

$$F(a) = \begin{cases} 0 & a \le \alpha \\ \frac{a-\alpha}{\beta-\alpha} & \alpha < a < \beta \\ 1 & a \ge \beta \end{cases} \tag{10}$$

Returning to the previous instance that random variable is uniformly distributed over the interval (0, 1), the expectation of uniform (weighted average) would be

$$E[u] = \int_0^1 u du \tag{11}$$

$$= \frac{1}{2} u^2 \Big|_0^1 \tag{12}$$

$$= \frac{1}{2} \tag{13}$$

And for the variance, which classify the spread

$$Var[u] = \int_0^1 \left(u - \frac{1}{2}\right)^2 du \tag{14}$$

$$= \frac{1}{3}\left(u - \frac{1}{2}\right)^3 \Big|_0^1 \tag{15}$$

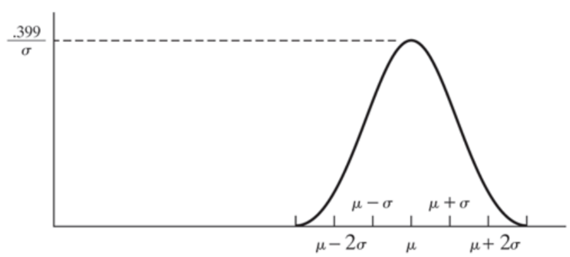
$$= \frac{1}{12} \tag{16}$$

2.2 Normal Distribution

We say that X is normally distributed, with parameters μ and σ^2 if the density of X is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} - \infty < x < \infty \tag{17}$$

This density function is a bell-shaped curve that is symmetric about μ. Here's a graph of normal density function with arbitrary μ and σ^2



(The maximum value of this curve will exist when $(x-\mu)^2$ on the exponential to be 0. As well as $(x-\mu)^2$ becomes non-zero, which means to be positive, the curve starts to decay, and eventually form a bell curve.)
In the density function of a normal random variable, μ controls the center of the distribution (since curve is symmetric about μ), and σ^2 controls the spread of distribution.

In order to prove that f(x) is indeed a probability density function, we need to show that

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx = 1 \tag{18}$$

to make the substitution $y = (x-\mu)/\sigma$, the above integral becomes

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2/2}dy \tag{19}$$

Thus, what we need to do is showing that

$$\int_{-\infty}^{\infty} e^{-y^2/2}dy = \sqrt{2\pi} \tag{20}$$

Let $I = \int_{-\infty}^{\infty} e^{-y^2/2}dy$, we get

$$\begin{align}
I^2 &= \int_{-\infty}^{\infty} e^{-y^2/2}dy \int_{-\infty}^{\infty} e^{-x^2/2}dx \tag{21}\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\left(y^2+x^2\right)/2}dydx \tag{22}\\
&= \int_{0}^{\infty}\int_{0}^{2\pi} e^{-r^2/2}rd\theta dr \tag{23}\\
&= 2\pi\int_{0}^{\infty} re^{-r^2/2}dr \tag{24}\\
&= -2\pi e^{-r^2/2}\Big|_{0}^{\infty} \tag{25}\\
&= 2\pi \tag{26}
\end{align}$$

Therefore, $I = \sqrt{2\pi}$ and the function $f(x)$ is proved.

Here's a sample of application.
Let S_0 stands for the price of stock today and S_1 stands for the price of stock tomorrow S_1 often modeled as

$$S_1 = S_0e^r \tag{27}$$

where r is the rate of return, modeled as a normal random variable

$$\frac{S_1}{S_0} = e^r \tag{28}$$

$$\ln\tfrac{S_1}{S_0} \sim \text{Normal}(\mu, \sigma^2)$$

Some Important Properties of Normal Distribution:

If $X \sim Normal(\mu, \sigma^2$, then $Y = aX + b \sim Normal(a\mu, a^2\sigma^2)$

In this case, when we look at the cumulative distribution function of Y, it can be found out

$$F_Y = P[Y \leq y] \tag{29}$$
$$= P\left[aX + b \leq y\right] \tag{30}$$
$$= P\left[x \leq \frac{y-b}{a}\right] \tag{31}$$
$$= F_X\left(\frac{y-b}{a}\right) \tag{32}$$

With the similar approach, we are able to get the density function for a $N(\mu a + b, a^2\mu^2)$ random variable

$$\frac{d}{dy}P\left[Y \leq y\right] = f\left(y\right) \tag{33}$$
$$= \frac{d}{dy}F_X\left(\frac{y-b}{a}\right) \tag{34}$$
$$= \frac{1}{a}f_X\left(\frac{y-b}{a}\right) \tag{35}$$
$$= \frac{1}{a}\frac{1}{\sqrt{2\pi}\sigma}e^{-\left(\left(\frac{y-b}{a}\right)-\mu\right)^2/2\sigma^2} \tag{36}$$
$$= \frac{1}{\sqrt{2\pi}a\sigma}e^{-(y-b-\mu a)^2/2\sigma^2a^2} \tag{37}$$

Hence, we could write the standardized form of the cumulative distribution function, which has no closed form solution

$$\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/2\sigma^2}dx$$

Since the linear transformation of normal distribution is still normal, we only need to do the simplest computation that $\mu = 0$ and $\sigma^2 = 1$, which is named Standard Normal Distribution.($Z \sim N(0,1)$)
Based on the parameters given, we can write the density function and cumulative distribution function of Standard Normal Distribution as

$$f(z) = \frac{1}{2\pi}e^{-\frac{z^2}{2}} \tag{38}$$
$$F(z) = \int_{-\infty}^{z} \frac{1}{2\pi}e^{-\frac{t^2}{2}}dt \tag{39}$$
$$= \Phi(z) \tag{40}$$

The linear transformation of Standard Normal Distribution $Y = aZ + b$ is also a normal distribution, in which $Y \sim N(b, a^2)$.

In order to obtain the cumulative distribution function of this normal distribution after linearly transforming, the equations can be written as

$$
\begin{aligned}
F(y) &= P\left[Y \leq y\right] && (41)\\
&= P\left[aZ + b \leq y\right] && (42)\\
&= P\left[Z \leq \frac{y-b}{a}\right] && (43)\\
&= \Phi\left(\frac{y-b}{a}\right) && (44)
\end{aligned}
$$

Hence, we already learn about the process that transforming from Standard Normal Distribution into an arbitrary normal distribution. Vice versa, if given a normal random variable $X \sim N(\mu, \sigma^2)$ then we are able to normalize it to a Standard Normal random variable $Z = \frac{X-\mu}{\sigma}$.

2.3 Binomial Approximation

Analogous to the Poisson approximation to binomial probabilities, which is a good approximation when n is large and p is small with $\lambda = np \sim o(1)$, a normal random variable is also an approach to approximate a binomial random variable where n is large, through DeMoivre-Laplace Limit Theorem

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

to standardize the random variable S_n

$$
\begin{aligned}
\lim_{n\to\infty} P\left[a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right] &= \Phi(b) - \Phi(a) && (45)\\
&= \int_a^b f(x)dx && (46)
\end{aligned}
$$

where $f(x)$ in the result stands for the standard normal density function

3 Synchronous Notes

3.1 Uniform Random Variable Recap

The uniform random variable is one that takes values on an interval with equal probability.

$U \sim uniform(\alpha, \beta)$ if

$P[U \in (a, b)] = \int_a^b \frac{1}{(\beta-\alpha)} dx$

where $(\alpha < a < b < \beta)$.

The PDF (probability density function) is a constant: $\frac{1}{(\beta-\alpha)}$

$$PDF : f(u) = \begin{cases} \frac{1}{(\beta-\alpha)} & \text{if } x \in (a,b) \\ 0 & \text{otherwise} \end{cases}$$

Expected Value:

$$E[U] = \tfrac{1}{2}(\alpha + \beta$$

Variance:

$$\begin{aligned} Var[U] &= E[U^2] - (E[U])^2 \\ &= \int_\alpha^\beta u^2 \frac{1}{\beta-\alpha} du - \frac{1}{4}(\alpha+\beta)^2 \\ &= \frac{1}{3}u^3 \frac{1}{(\beta-\alpha)}\Big|_\alpha^\beta - \frac{1}{4}(\alpha+\beta)^2 \\ &= \frac{1}{3(\beta-\alpha)}(\beta^3-\alpha^3) - \frac{1}{4}(\alpha+\beta)^2 \end{aligned}$$

Then, by factoring $(\beta^3 - \alpha^3)$ to $(\beta-\alpha)(\beta^2+\alpha\beta+\alpha^2)$, we see that the $(\beta-\alpha)$ terms cancel:

$$Var[U] = \tfrac{4(\alpha^2+\alpha\beta+\beta^2)-3(\alpha^2+2\alpha\beta+\beta^2)}{12}$$

Simplifying this expression yields:

$$Var[U] = \tfrac{(\alpha-\beta)^2}{12}$$

3.2 Normal Random Variable Recap

A **Normal Random Variable** has a PDF given by:

$$PDF : f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty < x < \infty$$

Expected Value:

$$E[X] = \mu$$

Variance:

$$Var[X] = \sigma^2$$

where μ is the value at which $f(x)$ is at a maximum and σ is the width of

the PDF.
If $X \sim N(\mu, \sigma^2)$, then

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

The cumulative distribution function for Z is

$$\begin{aligned} P[Z \leq z] &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx \\ &= \Phi(Z) \end{aligned}$$

This tells us that

$$\begin{aligned} P[Z \in (a,b]] &= P[Z \leq b] - P[Z \leq a] \\ &= \Phi(b) - \Phi(a) \end{aligned}$$

If we return to our random variable $X \sim N(\mu, \sigma^2)$, then

$$\begin{aligned} P[X \leq x] &= P[\frac{X-\mu}{\sigma} - \frac{x-\mu}{\sigma}] \\ &= P[Z \leq \frac{x-\mu}{\sigma}] \\ &= \Phi(\frac{x-\mu}{\sigma}) \end{aligned}$$

3.3 Student Questions:

Student Question 1: Is E[X] = x such that $P[X \leq x] = \frac{1}{2}$ for Normal Random Variables?

Answer: No. We know that the cumulative distribution function starts at 0 and increases to 1. This is generally only true if the PDF is symmetric about the expected value.

Counterexample:

$$\begin{aligned} f(x) &= 2x \\ F(x) &= \int_0^x 2tdt = x^2 \\ &\rightarrow F(x) = \frac{1}{2} \\ &\rightarrow x = \frac{1}{\sqrt{2}} \\ E[X] &= \int_0^1 x * 2xdx = \int_0^1 2x^2 dx = \frac{2}{3} \end{aligned}$$

If f is symmetric about μ, then $f(x-\mu)$ is symmetric about 0. Therefore, $\int xf(x-\mu)dx = 0$.

The new expected values are:

$$\rightarrow E[X-\mu]=0$$
$$\rightarrow E[X]=\mu$$

Student Question 2: Review normal approximation to the binomial distribution.
Answer: If $S_n \sim$ binomial(n, p), then

$$\frac{S_n - np}{\sqrt{np(1-p)}} \rightarrow N(0,1)$$
$$\text{as } n \rightarrow \infty$$

$$\begin{aligned}\lim_{n\rightarrow\infty} P[a < \frac{S_n - np}{\sqrt{np(1-p)}} \leq b] &= \Phi(b) - \Phi(a) \\ &= P[a < Z \leq b]\end{aligned}$$

where Z $\sim$ N(0,1).