PROBABILITY

QSNAKE EDITION

1. Review.

Definition.

- *Two events A and B are called mutually exclusive if $A \cap B = \emptyset$*
- *Events $A_1, A_2, A_3, \cdots$ are said to be mutually exclusive if they are pairwise mutually exclusive. That is, if $A_i \cap A_j = \emptyset$ whenever $i \neq j$.*

Theorem.

- $P(A) = 1 - P(A')$
- $P(A) \leq 1$, *for any event A*
- *For any two event A and B,* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- *If $A \subset B$, then* $P(A) \leq P(B)$
- $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$ *If $A_1, \cdots$ is a sequence of events*
- *If $A_1, A_2, \cdots, A_k$ are events, then* $P(\bigcap_{i=1}^{k} A_i) \geq 1 - \sum_{i=1}^{k} P(A_i')$

Proof. By axiom of probability. ■

Definition. *The conditional probability of an event A, given the event B, is defined by*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

if $P(B) \neq 0$

Theorem. *For any events A and B,*

$$P(A \cap B) = P(B)P(A \mid B) = P(A)P(B \mid A)$$

Theorem (Not talk in class)**.** *If $B_1, \cdots, B_k$ is a collection of mutually exclusive and exhaustive events, then for any event A*

$$P(A) = \sum_{i=1}^{k} P(B_i)P(A \mid B_i)$$

Note. exhaustive events: A collection of event which union is sample space.

Theorem (Not talk in class)**.** *If $B_1, \cdots, B_k$ is mutually exclusive and exhaustive events, then for any event A and each $j = 1, \cdots, k$*

$$P(B_j \mid A) = \frac{P(B_j)P(A \mid B_j)}{\sum_{i=1}^{k} P(B_i)P(A \mid B_i)} \left(= \frac{P(A \cap B_j)}{P(A)} \right)$$

Definition. *Two events A and B are called independent events if*

$$P(A \cap B) = P(A)P(B)$$

Otherwise, A and B are called dependent event

Theorem. *If A and B are events such that $P(A) > 0$ and $P(B) > 0$, and A and B are independent, we get*

$$P(A \cap B) = P(A)P(B) \Leftrightarrow P(A \mid B) = P(A) \Leftrightarrow P(B \mid A) = P(B)$$

Theorem (Not talk in class)**.**

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ \Leftrightarrow P(A' \cap B) &= P(A')P(B) \\ \Leftrightarrow P(A \cap B') &= P(A)P(B') \\ \Leftrightarrow P(A' \cap B') &= P(A')P(B') \end{aligned}$$

Definition. *The k events $A_1, \cdots, A_k$ are said to b independent or mutually independent if for every $j = 2, 3, \cdots, k$ and every subset of distinct indices $i_1, i_2, \cdots, i_j$*

$$P(A_{i1} \cap A_{i2} \cap \cdots \cap A_{ij}) = P(A_{i1})P(A_{i2}) \cdots P(A_{ij})$$

2. Discrete Random Variable Review.

Definition. *The cumulative distribution function(CDF) of a random variable X is defined for any real x by*

$$F(x) = P[X \leq x]$$

Theorem. *Let X be a discrete random variable with pdf $f(x)$ and CDF $F(x)$. If the possible values of X are indexed in increasing order, $x_1 < x_2 < x_3 < \cdots$, then:*
(i) $f(x_1) = F(x_i)$
(ii) for any $i > 1$, $f(x_i) = F(x_i) - F(x_{i-1})$
(iii) if $x < x_1$ *then* $F(x) = 0$
(iv) $F(x) = \sum_{x_i \leq x} f(x_i)$

Theorem. *A function $F(x)$ is a CDF for some random variable X if and only if it satisfies:*
(i) $\lim_{x \to -\infty} F(x) = 0$
(ii) $\lim_{x \to \infty} F(x) = 1$
(iii) $\lim_{h \to 0^+} F(x+h) = F(x)$
(iv) $a < b$ *implies* $F(a) \leq F(b)$

Example.
Suppose that the distribution function of X is given by

$$F(b) = \begin{cases} 0 & b < 0 \\ \frac{b}{4} & 0 \leq b < 1 \\ \frac{1}{2} + \frac{b-1}{4} & 1 \leq b < 2 \\ \frac{11}{12} & 2 \leq b < 3 \\ 1 & 3 \leq b \end{cases}$$

(1) Find $P\{X = i\},\ i = 1, 2, 3$
(2) Find $P\{\frac{1}{2} < X < \frac{3}{2}\}$

3. Continuous Random Variables.

Definition. *Let X be such a random variable. We say that X is a continuous random variable if there **exists a nonnegative function** f**, define for all real** $x \in (-\infty, \infty)$, having the property that for any set B of real numbers,*

$$P\{X \in B\} = \int_B f(x)dx$$

the function f is called the probability density function of the random variable X

Note. X will be in B may be obtained by integrating the probability density function over the set B. Since X must assume some value, f must satisfy

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx$$

All probability statements about X can be answered in terms of f, and letting $B = [a, b]$, we obtain

$$P\{a \leq X \leq b\} = \int_a^b f(x)dx$$

and we let $a = b$ we get

$$P\{X = a\} = \int_a^a f(x)dx = 0$$

this equation states that the probability that a continuous random variable will assume any fixed value is zero. Hence, for a continuous random variable,

$$P\{X < a\} = P\{X \leq a\} = F(a) = \int_{\infty}^{a} f(x)dx$$

Definition. *If X is a continuous random variable having probability density function $f(x)$, then expected value of X is*

$$E[X] = \int_{\infty}^{\infty} xf(x)dx$$

Proposition. If X is a continuous random variable with probability density function $f(x)$, then for any real-valued function g,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

To proof part of this proposition $(g(x) \geq 0)$, we will need the following lemma.(The general proof, which follows the argument in the case we present, is indicated in Theoretical Exercises 2 and 3.)

Lemma. *For a nonnegative random variable Y.*

$$E[Y] = \int_0^{\infty} P\{Y > y\}dy$$

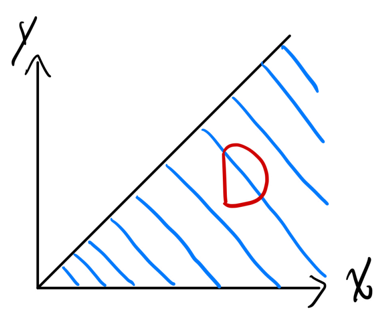
Proof. We present a proof when Y is a continuous random variable with probability density function f_Y. We have

$$\int_0^{\infty} P\{Y > y\}dy = \int_0^{\infty}\int_y^{\infty} f_Y(x)dxdy$$

where we have used the fact that $P\{Y > y\} = \int_y^{\infty} f_Y(x)dx$, and let us change it to region D first.

$$\int_0^{\infty}\int_y^{\infty} f_Y(x)dxdy = \iint_D \sin(y^2)dA$$

where $D = \{(x, y) \mid y < x < \infty,\ 0 < y < \infty\}$ the picture is



we can consider D as

$$D = \{(x, y) \mid 0 \leq x \leq \infty,\ 0 \leq y \leq x\}$$

and interchanging the order of integration in the preceding equation yields

$$
\begin{aligned}
\int_0^\infty P\{Y > y\}dy &= \iint_D F_Y(x)dA \\
&= \int_0^\infty \left(\int_0^x dy\right) f_Y(x)dx \\
&= \int_0^\infty x f_Y(x)dx \\
&= E[Y]
\end{aligned}
$$

■

Proof of Prop 2.1

From Lemma 2.1, for any function g for which $g(x) \geq 0$

$$
\begin{aligned}
E[g(X)] &= \int_0^\infty P\{g(X) > y\}dy \\
&= \int_0^\infty \int_{x:g(x)>y} f(x)dxdy \\
&= \int_{x:g(x)>0} \int_0^{g(x)} dy f(x)dx \\
&= \int_{x:g(x)>0} g(x)f(x)dx
\end{aligned}
$$

which completes the proof

Corollary. *If a and b are constants, then*

$$E[aX + b] = aE[X] + b$$

3. Uniform Random Variable.

Definition. *A random variable is said to be uniformly distributed over the interval* $(0, 1)$ *if its probability density function is given by*

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

and in general, we say that X *is a uniform random variable on the interval* (α, β) *if the probability density function of* X *is given by*

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & if\ \alpha < x < \beta \\ 0 & otherwise \end{cases}$$

We usy $X \sim Uniform(\alpha, \beta)$

Exercise. Let X be uniformly distributed over (α, β). Find $E[X]$ and $Var(X)$

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_{\alpha}^{\beta} \frac{x}{\beta - \alpha} dx \\ &= \frac{\beta^2 - \alpha^2}{2(\beta - \alpha)} \\ &= \frac{\beta + \alpha}{2} \end{aligned}$$

the variance be the homework, try it yourself.

4. Normal Random Variables.

Definition. *We say that X is a normal random variable, or simply that X is normally distributed, with parameters μ and σ^2 if the density of X is given by*

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

and we will usually write $X \sim N(\mu, \sigma^2)$

Proof $f(x)$ is indeed a probability density function

Proof. To prove that $f(x)$ is indeed a probability density function, we need to show that

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

Making the substitution $y = (x - \mu)/\sigma$, we see that

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

Hence, we must show that

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

Toward this end, let $I = \int_{-\infty}^{\infty} e^{-y^2/2} dy$. Then

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-y^2/2} dy \int_{-\infty}^{\infty} e^{-x^2/2} dx \\ &= 2\pi \int_0^{\infty} r e^{-r^2/2} dr \\ &= -2\pi e^{-r^2/2}|_0^{\infty} \\ &= 2\pi \end{aligned}$$

Hence, $I = \sqrt{2\pi}$, and the result is proved. ■

Note.
If X is normally distributed with parameters μ and σ^2, then $Y = aX + b$ is normally distributed with parameters $a\mu + b$ and $a^2\sigma^2$

Proof. Suppose $a > 0$. Let F_Y denote the cumulative distribution function of Y. Then

$$\begin{aligned} F_Y(x) &= P\{Y \leq x\} \\ &= P\{aX + b \leq x\} \\ &= P\left\{X \leq \frac{x-b}{a}\right\} \\ &= F_X\left(\frac{x-b}{a}\right) \end{aligned}$$

and by differentiation, the density function of Y is then

$$\begin{aligned} f_Y(x) &= \frac{1}{a} f_X\left(\frac{x-b}{a}\right) \\ &= \frac{1}{\sqrt{2\pi} a\sigma} exp\left\{-\left(\frac{x-b}{a} - \mu/2\sigma^2\right)\right\} \\ &= \frac{1}{\sqrt{2\pi} a\sigma} exp\{-(x-b-a\mu)^2/2(a\sigma)^2\} \end{aligned}$$

which shows that Y is normal with parameters $a\mu + b$ and $a^2\sigma^2$ ∎

Exercise. Find $E[X]$ and $Var(X)$ when X is a normal random variable with parameters μ and σ^2

Note. If X is normally distributed with parameters μ and σ^2, then $Z = (X - \mu)/\sigma$ is normally distributed with parameters 0 and 1. Such a random variable is said to be a standard, or a unit normal random variable.

Proof. Start by finding the mean and variance of the standard normal random variable $Z = (X - \mu)/\sigma/$ We have

$$\begin{aligned} E[Z] &= \int_{-\infty}^{\infty} x f_Z(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2} dx \\ &= -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} |_{-\infty}^{\infty} \\ &= 0 \end{aligned}$$

Thus,

$$\begin{aligned} Var(Z) &= E[Z^2] \\ &= \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} x^2 e^{-x^2/2} dx \end{aligned}$$

Integration by parts (with $u = x$ and $dv = xe^{-x^2/2}$) now gives

$$\begin{aligned} Var(Z) &= \frac{1}{\sqrt{2\pi}} \left(-e^{-x^2/2}|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2} dx \right) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} e^{-x^2/2} dx \\ &= 1 \end{aligned}$$

■

Because $X = \mu + \sigma Z$, the preceding yields the results

$$E[X] = \mu + \sigma E[Z] = \mu \text{ and } Var(X) = \sigma^2 Var(Z) = \sigma^2$$

Definition. *If $X \sim N(\mu, \sigma^2)$ then $X = \sigma Z + \mu$ where $Z \sim N(0, 1)$. We call Z a standard normal random variable.*

Definition. *It is customary to denote the cumulative distribution function of a standard normal random variable by $\Phi(x)$. That is ,*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$$

5. Exponential Random Variables.

<u>Definition</u>. *A continuous random variable whose probability density function is given, for some $\lambda > 0$, by*

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \textit{if } x \geq 0 \\ 0 & \textit{if } x < 0 \end{cases}$$

is said to be an exponential random variable with parameter λ. The cumulative distribution function $F(a)$ of an exponential random variable is given by

$$\begin{aligned} F(a) &= P\{X \leq a\} \\ &= \int_0^a \lambda e^{-\lambda x} dx \\ &= -e^{-\lambda x}|_0^a \\ &= 1 - e^{\lambda a} \qquad a \geq 0 \end{aligned}$$

Exercise. Let X be an exponential random variable with parameter λ. Calculate $E[X]$ an $Var(X)$.

Hazard Rate Functions

<u>Definition</u>. *Consider a positive continuous random variable X that we interpret as being the life time of some item. Let X have distribution function F and density f. The hazard rate (sometimes called the failure rate) function $\lambda(t)$ of F is defined by*

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)}, \qquad \textit{where } \overline{F} = 1 - F$$

Suppose now that the lifetime distribution is exponential. Then, by the memory-less property, it follows that the distribution of remaining life for a t-year-old item is the same as that for a new item.

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$$

6.* **The Gamma Distribution.** A random variable is said to have a gamma distribution with parameters $(\alpha, \lambda), \lambda > 0, \alpha > 0$, if its density function is given by

$$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where $\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy$ is called **gamma function**.
Note. Some times, the λ in the equation above will be changed to β, and normally $\beta = \dfrac{1}{\lambda}$
Now, let's check the property of $\Gamma(\alpha)$, integration of $\Gamma(\alpha)$ by parts yields

$$\begin{aligned} \Gamma(\alpha) &= -e^{-y} y^{\alpha-1}|_0^\infty + \int_0^\infty e^{-y}(\alpha-1)y^{\alpha-2} dy \\ &= (\alpha-1)\int_0^\infty e^{-y} y^{\alpha-2} dy \\ &= (\alpha-1)\gamma(\alpha-1) \end{aligned}$$

For integral values of α, say, $\alpha = n$, we obtain, by applying Equation above,

$$\begin{aligned} \Gamma(n) &= (n-1)\gamma(n-1) \\ &= (n-1)(n-2)\Gamma(n-2) \\ &= \cdots \\ &= (n-1)(n-2)\cdots 3 \cdot 2\Gamma(1) \end{aligned}$$

Since $\Gamma(1) = \int_0^\infty e^{-x} dx = 1$, it follows that, for integral values of n

$$\Gamma(n) = (n-1)!$$

Note.

(1) When $\alpha = 1$, this distribution reduces to the exponential distribution.
(2) When $\lambda = \frac{1}{2}$ and $\alpha = n/2$, n a positive integer, is called the χ_n^2(chi-squared) distribution with n degrees of freedom.

7. The Distribution of a Function of a Random Variable. Some times, we know the distribution of X and want to find the distribution of $g(X)$. To do so, it is necessary to express the event that $g(X) \leq y$ in terms of X being in some set.

Example. Let X be uniformly distributed over $(0, 1)$. We obtain the distribution of the random variable Y, defined by $Y = X^n$, as follows: For $0 \leq y \leq 1$,

$$
\begin{aligned}
F_Y(y) &= P\{Y \leq y\} \\
&= P\{X^n \leq y\} \\
&= P\{X \leq y^{1/n}\} \\
&= F_X(y^{1/n}) \\
&= y^{1/n}
\end{aligned}
$$

For instance, the density function of Y is given by

$$
f_Y(y) = \begin{cases} \frac{1}{n} y^{1/n-1} & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}
$$

Example. If X is a continuous random variable with probability density f_X, then the distribution of $Y = X^2$ is obtained as follows. For $y \geq 0$,

$$
\begin{aligned}
F_Y(y) &= P\{Y \leq y\} \\
&= P\{X^2 \leq y\} \\
&= P\{-\sqrt{y} \leq X \leq \sqrt{y}\} \\
&= F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned}
$$

Differentiation yields

$$
f_Y(y) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})]
$$

Theorem. *Let X be a continuous random variable having probability density function f_X. Suppose that $g(x)$ is a strictly monotonic (increasing or decreasing), differentiable (and thus continuous) function of x.*

Then the random variable Y defined by $Y = g(X)$ has a probability density function given by

$$f_Y(y) = \begin{cases} f_X[g^{-1}(y)] \left| \dfrac{d}{dy} g^{-1}(y) \right| & \text{if } y = g(x) \text{ for some } x \\ 0 & \text{if } y \neq g(x) \text{ for all } x \end{cases}$$

where $g^{-1}(y)$ is defined to equal that value of x such that $g(x) = y$.

Proof. Suppose that $y = g(x)$ for some x. Then, with $Y = g(X)$,

$$\begin{aligned} F_Y(y) &= P\{g(X) \leq y\} \\ &= P\{X \leq g^{-1}(y)\} \\ &= F_X(g^{-1}(y)) \end{aligned}$$

Differentiation gives

$$f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

■

Homework. The Lognormal Distribution (p.225)
If X is a normal random variable with mean μ and variance σ^2, then the random variable

$$Y = e^X$$

is said to be a lognormal random variable with parameters μ and σ^2. Try to find the density function f_Y