# Ve301 Recitation Class (Probability Section)

Wang Qian

UM-SJTU Joint Institute

2014-2015 Spring

# Personal Info.

Wang Qian

- Office hour: Friday 12:50 to 13:50 at e-reading room
- Recitation class: Friday 10:00 to 11:40 at F100
- e-mail: wangqian1992511@sjtu.edu.cn
- Tel: 18817519155
- Skype: wangqian1992511

When sending an e-mail to me, please begin your subject with "[VE301]", so that I can quickly recognize it. It can also ensure that your mail (especially the Tencent Mail) cannot be blocked. I may not be able to reply your short message at once, but if you have something urgent, please feel free to call me.

# Contents I


1. Introduction to Probability
2. Discrete Random Variables
3. Continuous Random Variables

# Question

Romeo wants to make a phone call to Juliet, but he forgets the last digit of Juliet's phone number. What is probability that Romeo can successfully express his love, "If I profane with my unworthiest hand. This holy shrine, the gentle sin is this: My lips, two blushing pilgrims, ready stand. To smooth that rough touch with a tender kiss", within three attempts?

# Answer

Let $A_i$ be the event that "the $i$-th attempt is successful", $A$ be the event that "Romeo successfully expressed his love to Juliet within three attempts" and $B$ be the event "the three attempt are all unsuccessful". Then, we observe that:

$$
\begin{aligned}
B &= \bar{A}_1\bar{A}_2\bar{A}_3 \\
A &= \bar{B} \\
P(A) &= 1 - P(B) \\
&= 1 - P(\bar{A}_1\bar{A}_2\bar{A}_3) \\
&= 1 - P(\bar{A}_3 \mid \bar{A}_1\bar{A}_2)P(\bar{A}_2 \mid \bar{A}_1)P(\bar{A}_1) \\
&= 1 - \frac{7}{8} \times \frac{8}{9} \times \frac{9}{10} \\
&= \frac{3}{10}
\end{aligned}
$$

# Question

Tom and Jerry are playing the dice. During first round, Tom throws the dice first and Jerry throws after Tom. The winner for every round is the one who gets the number larger than three first. The loser will always throw the dice first in the next coming round. What is the probability for Tom to win during the $n$-th round?

# Answer

Let $P_n$ be the probability that "Tom wins during the $n$-th round". Then, we observe that:

$$
\begin{aligned}
P_1 &= \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \ldots.. = \frac{2}{3} \\
P_{n+1} &= \frac{2}{3}(1 - P_n) + \frac{1}{3}P_n = -\frac{1}{3}P_n + \frac{2}{3}
\end{aligned}
$$

Introduce the parameter $\lambda$ such that:

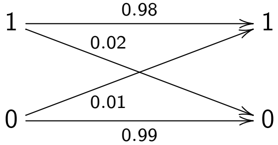
$$
\lambda = \frac{\frac{2}{3}}{-\frac{1}{3} - 1} = -\frac{1}{2}
$$

We will find that:

$$
\begin{aligned}
P_{n+1} - \frac{1}{2} &= -\frac{1}{3}(P_n - \frac{1}{2}) \\
P_n &= \frac{1}{2}[1 - (-\frac{1}{3})^n]
\end{aligned}
$$

# Question

Consider the following channel:




By statistics, the frequency of 1 and 0 is about 2:1. Suppose that one receives a code "1001", what is the probability that the original code transmitted is also "1001"?

# Answer

Let $T$ be the event that "transmitter sends bit 1", $R$ be the event that "the receiver receives bit 1", $A$ be the event that "the received code is the same as the transmitted code". Then we have:

$$P(A) = P(T|R)P(\bar{T}|\bar{R})P(\bar{T}|\bar{R})P(T|R)$$

By Bayes's Theorem, we can derive that:

$$P(T|R) = \frac{P(R|T)P(T)}{P(R|T)P(T) + P(R|\bar{T})P(\bar{T})} = \frac{196}{197}$$

$$P(\bar{T}|\bar{R}) = \frac{P(\bar{R}|\bar{T})P(\bar{T})}{P(\bar{R}|T)P(T) + P(\bar{R}|\bar{T})P(\bar{T})} = \frac{99}{103}$$

Thus,

$$P(A) = P(T|R)P(\bar{T}|\bar{R})P(\bar{T}|\bar{R})P(T|R) = 91.45\%$$

# Question

There are $m + n$ one cent coins in a package. Among them, $n$ coins are fake, which means that both of their faces are "Abraham Lincoln". If we pick up one coin from this package and toss it for $k$ times, we will observe that it shows "Abraham Lincoln" every time. What is the probability that this coin is not fake?

# Answer

Let $A$ be the event that "this coin is not fake", and $B$ be the event that "we will observe that it shows Abraham Lincoln every time during the $k$ tosses". By Bayes's Theorem,

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})} \\ &= \frac{\frac{1}{2^k} \times \frac{m}{m+n}}{\frac{1}{2^k} \times \frac{m}{m+n} + 1 \times \frac{n}{m+n}} \\ &= \frac{m}{m + 2^k n} \end{aligned}$$

# Question

Prove that the necessary and sufficient condition for event $A$ and $B$ to be independent is that $P(A|B) = P(A|\bar{B})$.

# Answer

To prove the necessity, we suppose that $A$ and $B$ are independent, when $A$ and $\bar{B}$ are independent as well. Thus,

$$P(A|B) = P(A) = P(A|\bar{B})$$

To prove the sufficiency, we suppose that $P(A|B) = P(A|\bar{B})$. By definition,

$$\frac{P(AB)}{P(B)} = \frac{P(A\bar{B})}{P(\bar{B})}$$

From the property of ratio, we derive that:

$$\frac{P(AB)}{P(B)} = \frac{P(AB) + P(A\bar{B})}{P(B) + P(\bar{B})} = \frac{P(A(B \cup \bar{B}))}{1} = P(A)$$

This means that $P(AB) = P(A)P(B)$, proving the independency of event $A$ and $B$.

# Question

Given a hash function $H$, with $n$ possible outputs and a specific value $H(x)$, if $H$ is applied to $k$ random inputs, what must be the value of $k$ so that the probability of an arbitrary hash collision among the $k$ outputs would be 0.5?

# Answer

The solution is based on Birthday Paradox. The probability for all these $k$ outputs are different is:

$$
\begin{aligned}
P &= \frac{n}{n} \times \frac{n-1}{n} \times \frac{n-2}{n} \times \ldots \times \frac{n-(k-1)}{n} \\
&= (1-\frac{1}{n})(1-\frac{2}{n})\ldots(1-\frac{k-1}{n}) \\
&\approx e^{\frac{1}{n}} e^{\frac{2}{n}} \ldots e^{\frac{k-1}{n}} \\
&= e^{-\frac{k(k-1)}{2n}}
\end{aligned}
$$

Thus, we can conclude that $k = 1.2\sqrt{n}$.

# Question

Consider the extended Monty Hall Paradox. Suppose that we have $n$ doors. Behind one of the doors is the prize, behind the other $n-1$ doors there is nothing. Here, $n = 2m+1$ when $m$ is a positive integer. The rules for the game are as following:

1. Choose one of the closed and not yet chosen doors.
2. If no more closed and not yet chosen doors exist, the game is over.
3. Game master opens one of the closed and not yet chosen empty doors.
4. Return to step one or the game is over.

What is probability that the door with the prize behind is chosen in the end, if the "always exchanging" strategy is used?

# Answer

When $n = 3$, we know that $P(3) = \frac{2}{3}$.
Now, we should derive the relation between $P(k)$ and $P(k+2)$ by induction. Notice that, we can have the chance to win the prize and reduce the number of empty doors by two, if and only if we choose an empty door at first under this circumstance. Therefore,

$$P(k+2) = \frac{k+1}{k+2} P(k)$$

Then, we can derive that:

$$P(n) = \frac{2 \times 4 \times 6 \times ... \times (n-1)}{3 \times 5 \times 7 \times ... \times n}$$

# Question

There is one black ball and one white ball in a box. Catch a ball randomly from the box. If the ball is white, the game is over. Otherwise, put the ball back to the box, and put back one more ball into the box. Does the exception of the rounds of the game, $E[X]$, exist, if the added ball is always white? What if the added ball is always black?

# Answer

Let's first discuss the condition when the added ball is always white.

$$\begin{aligned}
E[X] &= 1 \times \frac{1}{2} + 2 \times \frac{1}{2} \times \frac{2}{3} + 3 \times \frac{1}{2} \times \frac{1}{3} \times \frac{3}{4} + 4 \times \frac{1}{2} \times \frac{1}{3} \times \frac{1}{4} \times \frac{4}{5} + \dots \\
&< 2 \times \frac{1}{2} + 3 \times \frac{1}{2} \times \frac{2}{3} + 4 \times \frac{1}{2} \times \frac{1}{3} \times \frac{3}{4} + 5 \times \frac{1}{2} \times \frac{1}{3} \times \frac{1}{4} \times \frac{4}{5} + \dots \\
&= \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots \\
&= e
\end{aligned}$$

Therefore, $E[X]$ exists.

# Answer

Now, the added ball is always black.

$$
\begin{aligned}
E[X] &= 1\times\frac{1}{2} + 2\times\frac{1}{2}\times\frac{1}{3} + 3\times\frac{1}{2}\times\frac{2}{3}\times\frac{1}{4} + 4\times\frac{1}{2}\times\frac{2}{3}\times\frac{3}{4}\times\frac{1}{5} + \dots \\
&= \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots \\
&= \infty
\end{aligned}
$$

Therefore, $E[X]$ does not exist.

# Question

$X$ is a random variable. $C$ is a constant. Prove that $Var[X] \leq E[(X - C)^2]$.

# Answer

$$\begin{aligned} E[(X-C)^2] &= E[X^2 - 2CX + C^2] \\ &= E[X^2] - 2CE[X] + C^2 \\ &= (E[X^2] - E^2[X]) + (E^2[X] - 2CE[X] + C^2) \\ &= Var[X] + (E[X] - C)^2 \\ &\geq Var[X] \end{aligned}$$

The equality can be obtained when $C = E[X]$.

# Question

Throw a dice until all the six numbers appear at least once. Calculate the exception $E[X]$ of the rounds of throwing, $X$.

# Answer

Introduce random variable $Y_i$, where $i = 0, 1, 2, 3, 4, 5$:

- $Y_0 = 1$
- $Y_i$ is the rounds to get the $(i+1)$-th number after obtaining $i$ different numbers

Thus, it is obvious that:

$$X = Y_0 + Y_1 + Y_2 + Y_3 + Y_4 + Y_5$$

Notice that $Y_i$ follows a geometric distribution with parameter $\frac{6-i}{6}$. Therefore,

$$E[Y_i] = \frac{6}{6 - i}$$

Now, we can conclude that:

$$E[X] = E[Y_0] + E[Y_1] + E[Y_2] + E[Y_3] + E[Y_4] + E[Y_5] = 14.7$$

# Question

Two snipers Hathcock and Waldron are shooting in turns until they make two successful shoots. Given that the probability of a successful shoot is $p$, and the shoots are independent of each other, what is the probability that these two shoots are made by the same sniper?

# Answer

This is a problem about Pascal distribution. Let $n$ to be an arbitrary positive integer. The probability for them to end shooting at the $[2(n+1)]$-th attempt, while satisfying this condition, is:

$$P_0(n) = (1-p)^{n+1} C_n^{2-1} p^2 (1-p)^{n-1} = np^2(1-p)^{2n}$$

The probability for them to end shooting at the $(2n+1)$-th attempt, while satisfying this condition, is:

$$P_1(n) = (1-p)^{n} C_n^{2-1} p^2 (1-p)^{n-1} = np^2(1-p)^{2n-1}$$

Hence, the probability that the two shoots are made by the same sniper is:

$$\begin{aligned} P &= \sum_{n=1}^{\infty} P_0(n) + \sum_{n=1}^{\infty} P_1(n) \\ &= p^2(1-p)(2-p) \sum_{n=1}^{\infty} n[(1-p)^2]^{n-1} \end{aligned}$$

# Answer

Notice that $(1-p)^2$ is always smaller than one, therefore:

$$\begin{aligned} P &= p^2(1-p)(2-p)\sum_{n=1}^{\infty} n[(1-p)^2]^{n-1} \\ &= \frac{p^2(1-p)(2-p)}{[1-(1-p)^2]^2} \\ &= \frac{1-p}{2-p} \end{aligned}$$

# Question

Suppose that there are 1000 cars passing through a channel per day. The probability for a car to break down is 0.0001. What is the probability that there are no less than two cars breaking down per day?

# Answer

Let random variable $X$ to be the number of cars breaking down per day in this channel. We know that $X$ follows the binomial distribution with parameter $n = 1000$ and $p = 0.0001$. Thus, it is possible to approximate the binomial distribution with the Poisson distribution with parameter $k = np = 0.1$ in this situation.

$$\begin{aligned} P(X = x) &= C_n^x p^x (1-p)^{n-x} \\ &\approx \frac{k^x e^{-k}}{x!} \end{aligned}$$

Therefore, we can conclude that:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &\approx 1 - e^{-0.1} - e^{-0.1} * 0.1 \\ &= 0.0047 \end{aligned}$$

# Question

$X$ follows Poisson distribution with parameter $\lambda$. Find $E[\frac{1}{1+X}]$.

# Answer

$$
\begin{aligned}
E[\frac{1}{1+X}] &= \sum_{k=0}^{\infty} \frac{P[X=k]}{1+k} \\
&= \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k+1)!} \\
&= \frac{e^{-\lambda}}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{(k+1)!} \\
&= \frac{e^{-\lambda}}{\lambda} (\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} - 1) \\
&= \frac{e^{-\lambda}}{\lambda} (e^{\lambda} - 1) \\
&= \frac{1 - e^{-\lambda}}{\lambda}
\end{aligned}
$$

# Question

The cumulative distribution of the random variable $X$ is:

$$F(x) = \begin{cases} 0 & (x < 1) \\ ln(x) & (1 \leq x \leq e) \\ 1 & (x \geq e) \end{cases}$$

Find its corresponding density function.

# Answer

Since the density function is the derivative of the cumulative distribution, we have:

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

# Question

The density function of Maxwell distribution is

$$f(x) = \begin{cases} Ax^2 exp(-x^2/b) & (x > 0) \\ 0 & (otherwise) \end{cases}$$

Given that $b = m/(2kT)$, find the value of $A$.

# Answer

Take the integral of the density function:

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)dx &= \int_0^{\infty} Ax^2 exp(-x^2/b)dx \\
&= -\frac{Ab}{2} xexp(-x^2/b) \mid_0^{\infty} + \frac{Ab}{2}\int_0^{\infty} exp(-x^2/b)dx \\
&= \frac{Ab\sqrt{b}}{2}\int_0^{\infty} exp(-u^2)du \\
&= \frac{Ab\sqrt{b}}{4}\sqrt{\pi}
\end{aligned}$$

Remember the following useful conclusions:

$$\begin{aligned}
\int u^n exp(au)du &= \frac{1}{a}u^n exp(au) - \frac{n}{a}\int u^{n-1} exp(au)du \\
\int_{-\infty}^{\infty} exp(-t^2/2)dt &= \sqrt{2\pi}
\end{aligned}$$

# Question

A traffic light is in red for about 80% of the time. Someone is waiting for the green light. The waiting time $X$ is in the range of [0,30] seconds and follows the uniform distribution. Is $X$ a continuous random variable or a discrete one?

# Answer

Let $A$ be the event that the traffic light is green. For a certain $x \geq 0$, we have $(0 \leq x \leq 30)$:

$$\begin{aligned} P(X \leq x) &= P(X \leq x|A)P(A) + P(X \leq x|\bar{A})P(\bar{A}) \\ &= 1 \times 0.2 + \frac{x}{30} \times 0.8 \\ &= 0.2 + \frac{0.8x}{30} \end{aligned}$$

Thus,

$$F(x) = P(X \leq x) = \begin{cases} 0 & (x < 0) \\ 0.2 + \frac{0.8x}{30} & (0 \leq x \leq 30) \\ 1 & (x > 30) \end{cases}$$

Hence, this random variable is neither continuous nor discrete.

# Question

In one circuit, the voltage of a resistor follows the distribution $N(120, 2^2)$. We test the voltage of this resistor for 5 times. Find the probability that two of these five values will not be in the range of [118,122].

# Answer

Let $X_i$ be the value of the voltage for the $i$-th measurement.

$$\begin{aligned} P(188 \le X_i \le 122) &= \Phi(\frac{122-120}{2}) - \Phi(\frac{188-120}{2}) \\ &= \Phi(1) - \Phi(-1) \\ &= 0.6826 \end{aligned}$$

Thus, from the binomial contribution, we have:

$$P = C_5^2 \times 0.31747^2 \times 0.6826^3 = 0.3204$$

# Question

The random variable $X$ follows the normal distribution $N(0,1)$. Calculate the density function of $Y = exp(X)$.

# Answer

We only need to consider the case when $y$ is larger than 0.

$$\begin{aligned} F(y) &= P(0 < Y \leq y) \\ &= P(0 < exp(X) \leq y) \\ &= P(X \leq ln(y)) \\ &= \Phi(ln(y)) \end{aligned}$$

Therefore,

$$f(y) = \frac{d}{dy}F(y) = \frac{d}{dx}\Phi(x)|_{x=ln(y)}\frac{1}{y} = \frac{1}{\sqrt{2\pi}}exp[-\frac{1}{2}ln^2(y)]\frac{1}{y}$$

# Question

The density function of the random variable $X$ is

$$f(x) = \begin{cases} 2x/\pi^2 & (0 < x < \pi) \\ 0 & (otherwise) \end{cases}$$

Find the density function of $Y = sin(X)$.

# Answer

Obviously, we have $0 \leq y \leq 1$.

$$\begin{aligned} F(y) &= P(0 \leq Y \leq y) \\ &= P(0 \leq sin(X) \leq y) \\ &= P(0 \leq X \leq asin(y)) + P(\pi - asin(y) \leq \pi) \\ &= \int_0^{asin(y)} \frac{2x}{\pi^2} dx + \int_{\pi - asin(y)}^{\pi} \frac{2x}{\pi^2} dx \\ &= \frac{2asin(y)}{\pi} \end{aligned}$$

Thus, when $0 < y < 1$:

$$f(y) = \frac{d}{dx} F(y) = \frac{2}{\pi\sqrt{1 - y^2}}$$

# Question

We want to use the random variable $X$ to judge the quality of one product. It has the density function

$$f(x) = \begin{cases} xexp(-x^2/2) & (x > 0) \\ 0 & (otherwise) \end{cases}$$

The error $Y$ follows the uniform distribution $U(-\epsilon, \epsilon)$. Find the density function for $Z = X + Y$.

# Answer

The density of $Z$ can be expressed by the convolution:

$$f_Z(z) = f_X * f_y = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx$$

Determine the domain:

$$\begin{cases} x > 0 \\ -\epsilon < z - x < \epsilon \end{cases} \Rightarrow \begin{cases} x > 0 \\ z - \epsilon < x < z + \epsilon \end{cases}$$

Thus,

$$\begin{aligned} f_Z(z) &= \begin{cases} \frac{1}{2\epsilon}\int_0^{z+\epsilon} xexp(-x^2/2)dx & (-\epsilon < z < \epsilon) \\ \frac{1}{2\epsilon}\int_{z-\epsilon}^{z+\epsilon} xexp(-x^2/2)dx & (z \geq \epsilon) \\ 0 & (z \leq \epsilon) \end{cases} \\ &= \begin{cases} [1 - exp(-(x+z)^2/2)]/2\epsilon & (-\epsilon < z < \epsilon) \\ [exp(-(x-z)^2/2) - exp(-(x+z)^2/2)]/2\epsilon & (z \geq \epsilon) \\ 0 & (z \leq -\epsilon) \end{cases} \end{aligned}$$

# Question

Let $X$ and $Y$ to be independent random variables. They follow the exponential distribution with parameters $\alpha$ and $\beta$, respectively. Find $E[X^2 + Yexp(-X)]$

# Answer

$$
\begin{aligned}
E[X^2 + Yexp(-X)] &= E[X^2] + E[Y]E[exp(-X)] \\
&= Var[X] + E^2[x] + E[Y]m_X(-1) \\
&= \frac{1}{\alpha^2} + \frac{1}{\alpha^2} + \frac{1}{\beta}(1 + \frac{1}{\alpha})^{-1} \\
&= \frac{2}{\alpha^2} + \frac{\alpha}{\beta(\alpha + 1)}
\end{aligned}
$$

# Question

Random variables $X$ and $Y$ follow the standard normal distribution. They are independent from each other. Find $E[\frac{X^2}{X^2+Y^2}]$.

# Answer

From the symmetry, we know that:

$$E[\frac{X^2}{X^2+Y^2}] = E[\frac{Y^2}{X^2+Y^2}]$$

Besides, we also have the fact that:

$$E[\frac{X^2}{X^2+Y^2}] + E[\frac{Y^2}{X^2+Y^2}] = E[1] = 1$$

Therefore,

$$E[\frac{X^2}{X^2+Y^2}] = 0.5$$

# Question

The target point of an air-drop task is the origin point. The actual air-drop point is $(X, Y)$. Random variables $X$ and $Y$ follow the normal distribution $N(0, \sigma^2)$, and they are independent from each other. Let $D$ to be the distance between $(X, Y)$ and the origin point. Find $E[D]$.

# Answer

Since $X$ and $Y$ are independent, we derive that:

$$f(x,y) = \frac{exp(-(x^2+y^2)/2\sigma^2)}{2\pi\sigma^2}$$

Hence,

$$\begin{aligned}
E[D] &= E[\sqrt{X^2+Y^2}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sqrt{x^2+y^2}f(x,y)dxdy \\
&= 2\pi\int_0^{\infty}\frac{r^2 exp(-r^2/2\sigma^2)}{2\pi\sigma^2}dr = \int_0^{\infty}\frac{r^2 exp(-r^2/2\sigma^2)}{\sigma^2}dr \\
&= -\int_0^{\infty} rd(exp(-r^2/2\sigma^2)) = -rexp(-r^2/2\sigma^2)|_0^{\infty} + \int_0^{\infty} exp(-r^2/2\sigma^2)dr \\
&= \sigma\sqrt{\pi/2}
\end{aligned}$$

# Question

The density for the random vector $(X, Y)$ is

$$f(x,y) = \begin{cases} \frac{(x+y)}{2} exp(-(x+y)) & (x > 0, y > 0) \\ 0 & (otherwise) \end{cases}$$

Find if $X$ and $Y$ are independent or not.

# Answer

From the relation:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

We can derive that:

$$f_X(x) = \begin{cases} \frac{x+1}{2} exp(-x) & (x > 0) \\ 0 & (otherwise) \end{cases}$$

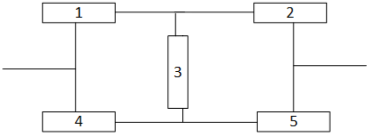
Similarly

$$f_Y(y) = \begin{cases} \frac{y+1}{2} exp(-y) & (y > 0) \\ 0 & (otherwise) \end{cases}$$

Since $f_X(x)f_Y(y) \neq f(x,y)$, $X$ and $Y$ are not independent.

# Question

Provided that the reliability of each component is $p$, calculate the reliability for the following system.



# Answer

Let $A$ be the event that "the whole system is reliable", and $A_i$ be the event that "component $i$ is reliable". From the total probability according to the status of component 3, we derive that:

$$P(A) = P(A|A_3)P(A_3) + P(A|\bar{A}_3)P(\bar{A}_3)$$

The remaining work is to calculate the reliability of a series parallel system and a parallel series system.

$$\begin{aligned} P(A) &= [1-(1-p)^2]^2 p + [1-(1-p^2)^2](1-p) \\ &= 2p^2 + 2p^3 - 5p^4 + 2p^5 \end{aligned}$$