

1 Predictable Gaussians

Note 21

Let Y be the result of a fair coin flip, and X be a normally distributed random variable with parameters dependent on Y . That is, if $Y = 1$, then $X \sim N(\mu_1, \sigma_1^2)$, and if $Y = 0$, then $X \sim N(\mu_0, \sigma_0^2)$.

(a) Sketch the two distributions of X overlaid on the same graph for the following cases:

(i) $\sigma_0^2 = \sigma_1^2, \mu_0 \neq \mu_1$

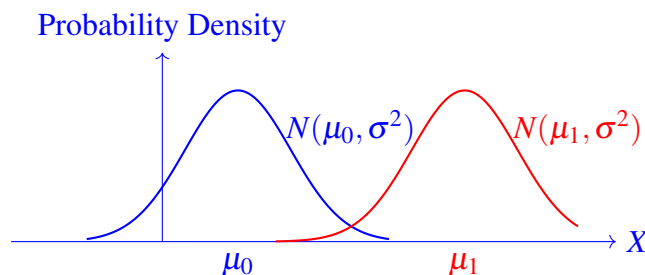
(ii) $\sigma_0^2 \neq \sigma_1^2, \mu_0 = \mu_1$

(b) Bayes' rule for mixed distributions can be formulated as $\mathbb{P}[Y = 1 \mid X = x] = \frac{\mathbb{P}[Y=1]f_{X|Y=1}(x)}{f_X(x)}$ where Y is a discrete distribution and X is a continuous distribution. Compute $\mathbb{P}[Y = 1 \mid X = x]$, and show that this can be expressed in the form of $\frac{1}{1+e^\gamma}$ for some expression γ . (Hint: any value z can be equivalently expressed as $e^{\ln(z)}$)

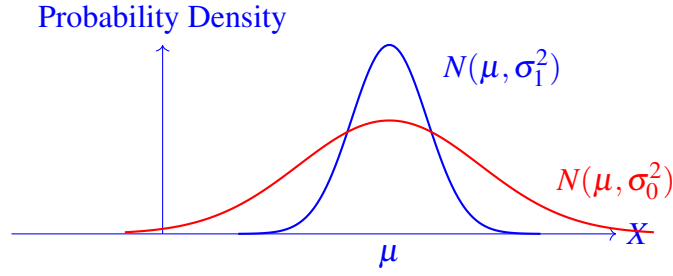
(c) In the special case where $\sigma_0^2 = \sigma_1^2$ find a simple expression for the value of x where $\mathbb{P}[Y = 1 \mid X = x] = \mathbb{P}[Y = 0 \mid X = x] = 1/2$, and interpret what the expression represents. (Hint: the identity $(a+b)(a-b) = a^2 - b^2$ may be useful)

Solution:

(a) (i) In this case, there are two bell curves with the same spread/width due to the variances being equal, but being centered at different means.



(ii) In this case, there will be two bell curves centered at the same mean, but the one with lower variance will be skinnier and taller, due to more of the probability density being centered closer to the mean.



(b)

$$\begin{aligned}
 & \mathbb{P}[Y = 1 \mid X = x] \\
 &= \frac{\mathbb{P}[Y = 1]f_{X|Y=1}(x)}{\mathbb{P}[Y = 1]f_{X|Y=1}(x) + \mathbb{P}[Y = 0]f_{X|Y=0}(x)} \\
 &= \frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right)}{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) + \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(x-\mu_0)^2}{2\sigma_0^2}\right)} \\
 &= \frac{1}{1 + \frac{\sigma_1}{\sigma_0} \exp\left(\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(x-\mu_0)^2}{2\sigma_0^2}\right)} \\
 &= \frac{1}{1 + \exp\left(\ln\left(\frac{\sigma_1}{\sigma_0}\right) + \frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(x-\mu_0)^2}{2\sigma_0^2}\right)}.
 \end{aligned}$$

Which is of the desired form, with $\gamma = \ln\left(\frac{\sigma_1}{\sigma_0}\right) + \left(\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(x-\mu_0)^2}{2\sigma_0^2}\right)$

(c) Note that $\mathbb{P}[Y = 1 \mid X = x] = \frac{1}{2}$ implies that $\exp(\gamma) = 1$, which means that $\gamma = 0$. Thus, $\ln\left(\frac{\sigma_1}{\sigma_0}\right) + \left(\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(x-\mu_0)^2}{2\sigma_0^2}\right) = 0$. Using the conditions from the problem statement, we can simplify this expression.

$$\begin{aligned}
 & \ln\left(\frac{\sigma_1}{\sigma_0}\right) + \left(\frac{(x-\mu_1)^2}{2\sigma^2} - \frac{(x-\mu_0)^2}{2\sigma^2}\right) = 0 \\
 & 0 + \left(\frac{(x-\mu_1)^2}{2\sigma^2} - \frac{(x-\mu_0)^2}{2\sigma^2}\right) = 0 \\
 & (x-\mu_1)^2 = (x-\mu_2)^2 \\
 & x^2 - 2\mu_1x + \mu_1^2 = x^2 - 2\mu_2x + \mu_2^2 \\
 & 2(\mu_2 - \mu_1)x = \mu_2^2 - \mu_1^2 \\
 & x = \frac{\mu_2^2 - \mu_1^2}{2(\mu_2 - \mu_1)} = \frac{\mu_2 + \mu_1}{2}
 \end{aligned}$$

Notice that x becomes the average, or center, of the two means.

2 Moments of the Gaussian

Note 21

For a random variable X , the quantity $\mathbb{E}[X^k]$ for $k \in \mathbb{N}$ is called the *kth moment* of the distribution. In this problem, we will calculate the moments of a standard normal distribution.

(a) Prove the identity

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{tx^2}{2}\right) dx = t^{-1/2}$$

for $t > 0$.

Hint: Consider a normal distribution with variance $\frac{1}{t}$ and mean 0.

(b) For the rest of the problem, X is a standard normal distribution (with mean 0 and variance 1). Use part (a) to compute $\mathbb{E}[X^{2k}]$ for $k \in \mathbb{N}$.

Hint: Try differentiating both sides with respect to t , k times. You may use the fact that we can differentiate under the integral without proof.

(c) Compute $\mathbb{E}[X^{2k+1}]$ for $k \in \mathbb{N}$.

Solution:

(a) Note that a normal distribution with mean 0 and variance t^{-1} has the density function

$$f(x) = \frac{\sqrt{t}}{\sqrt{2\pi}} \exp\left(-\frac{tx^2}{2}\right),$$

and since the density must integrate to 1, we see that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{tx^2}{2}\right) dx = t^{-1/2}.$$

(b) Differentiating the identity from (a) k times with respect to t , we obtain a LHS of

$$\begin{aligned} \frac{d^k}{dt^k} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{tx^2}{2}\right) dx \right] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{d^k}{dt^k} \left[\exp\left(-\frac{tx^2}{2}\right) \right] dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-1)^k \frac{x^{2k}}{2^k} \exp\left(-\frac{tx^2}{2}\right) dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{(-1)^k}{2^k} \int_{-\infty}^{\infty} x^{2k} \exp\left(-\frac{tx^2}{2}\right) dx \end{aligned}$$

Here, we use the fact that everything involving x is a constant with respect to t .

Looking at the RHS, we have

$$\frac{d^k}{dt^k} [t^{-1/2}] = (-1)^k \frac{1 \cdot 3 \cdots (2k-3) \cdot (2k-1)}{2^k} t^{-(2k+1)/2}$$

Together, this means that

$$\frac{1}{\sqrt{2\pi}} \frac{(-1)^k}{2^k} \int_{-\infty}^{\infty} x^{2k} \exp\left(-\frac{tx^2}{2}\right) dx = (-1)^k \frac{1 \cdot 3 \cdots (2k-3) \cdot (2k-1)}{2^k} t^{-(2k+1)/2}$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{2k} \exp\left(-\frac{tx^2}{2}\right) dx = (1 \cdot 3 \cdots (2k-3) \cdot (2k-1)) t^{-(2k+1)/2}$$

If we set $t = 1$, we get

$$\mathbb{E}[X^{2k}] = \int_{-\infty}^{\infty} x^{2k} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = \prod_{i=1}^k (2i-1).$$

This is sometimes denoted $(2k-1)!!$. Note that we can also write the result as

$$\mathbb{E}[X^{2k}] = (2k-1)!! = \frac{(2k)!}{2 \cdot 4 \cdots (2k-2) \cdot (2k)} = \frac{(2k)!}{2^k k!}.$$

(c) $\mathbb{E}[X^{2k+1}] = 0$, since the density function is symmetric around 0.

3 Chebyshev's Inequality vs. Central Limit Theorem

Note 17
Note 21

Let n be a positive integer. Let X_1, X_2, \dots, X_n be i.i.d. random variables with the following distribution:

$$\mathbb{P}[X_i = -1] = \frac{1}{12}; \quad \mathbb{P}[X_i = 1] = \frac{9}{12}; \quad \mathbb{P}[X_i = 2] = \frac{2}{12}.$$

(a) Calculate the expectations and variances of X_1 , $\sum_{i=1}^n X_i$, $\sum_{i=1}^n (X_i - \mathbb{E}[X_i])$, and

$$Z_n = \frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{n/2}}.$$

(b) Use Chebyshev's Inequality to find an upper bound b for $\mathbb{P}[|Z_n| \geq 2]$.

(c) Use b from the previous part to bound $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$.

(d) As $n \rightarrow \infty$, what is the distribution of Z_n ?

(e) We know that if $Z \sim \mathcal{N}(0, 1)$, then $\mathbb{P}[|Z| \leq 2] = \Phi(2) - \Phi(-2) \approx 0.9545$. As $n \rightarrow \infty$, provide approximations for $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$.

Solution:

(a) Firstly, let us calculate $\mathbb{E}[X_1]$ and $\text{Var}(X_1)$; we have

$$\mathbb{E}[X_1] = -\frac{1}{12} + \frac{9}{12} + \frac{4}{12} = 1$$

$$\text{Var}(X_1) = \frac{1}{12} \cdot 2^2 + \frac{9}{12} \cdot 0^2 + \frac{2}{12} \cdot 1^2 = \frac{1}{2}.$$

Using linearity of expectation and variance (since X_1, \dots, X_n are independent), we find that

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^n X_i\right] &= \sum_{i=1}^n \mathbb{E}[X_i] = n \\ \text{Var}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n \text{Var}(X_i) = \frac{n}{2}\end{aligned}$$

Again, by linearity of expectation,

$$\mathbb{E}\left[\sum_{i=1}^n (X_i - \mathbb{E}[X_i])\right] = \mathbb{E}\left[\sum_{i=1}^n X_i - n\right] = n - n = 0.$$

Subtracting a constant does not change the variance, so

$$\text{Var}\left(\sum_{i=1}^n (X_i - \mathbb{E}[X_i])\right) = \text{Var}\left(\sum_{i=1}^n X_i - n\right) = \frac{n}{2},$$

as before.

Using the scaling properties of the expectation and variance, we finally have

$$\begin{aligned}\mathbb{E}[Z_n] &= \mathbb{E}\left[\frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{n/2}}\right] = \frac{1}{\sqrt{n/2}} \mathbb{E}\left[\sum_{i=1}^n (X_i - \mathbb{E}[X_i])\right] = \frac{0}{\sqrt{n/2}} = 0 \\ \text{Var}(Z_n) &= \text{Var}\left(\frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{n/2}}\right) = \frac{1}{n/2} \text{Var}\left(\sum_{i=1}^n (X_i - \mathbb{E}[X_i])\right) = \frac{n/2}{n/2} = 1\end{aligned}$$

(b) Using Chebyshev's, we have

$$\mathbb{P}[|Z_n| \geq 2] \leq \frac{\text{Var}(Z_n)}{2^2} = \frac{1}{4}$$

since $\mathbb{E}[Z_n] = 0$ and $\text{Var}(Z_n) = 1$ as we computed in the previous part.

(c) $\frac{1}{4}$ for both, since we have

$$\begin{aligned}\mathbb{P}[Z_n \geq 2] &\leq \mathbb{P}[|Z_n| \geq 2] \\ \mathbb{P}[Z_n \leq -2] &\leq \mathbb{P}[|Z_n| \geq 2]\end{aligned}$$

(d) By the Central Limit Theorem, we know that $Z_n \rightarrow \mathcal{N}(0, 1)$, the standard normal distribution.

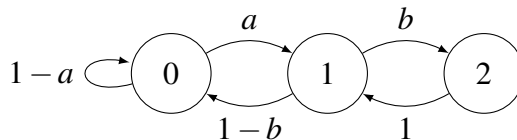
(e) Since $Z_n \rightarrow \mathcal{N}(0, 1)$, we can approximate $\mathbb{P}[|Z_n| \geq 2] \approx 1 - 0.9545 = 0.0455$. By the symmetry of the normal distribution, $\mathbb{P}[Z_n \geq 2] = \mathbb{P}[Z_n \leq -2] \approx 0.0455/2 = 0.02275$.

It is interesting to note that the CLT provides a much smaller answer than Chebyshev. This is due to the fact that the CLT is applied to a particular kind of random variable, namely the (scaled) sum of a bunch of random variables. Chebyshev's inequality, however, holds for any random variable, and is therefore weaker.

4 Analyze a Markov Chain

Note 22

Consider a Markov chain with the state diagram shown below where $a, b \in (0, 1)$.



Here, we let $X(n)$ denote the state at time n .

- Is this Markov chain irreducible? Is this Markov chain aperiodic? Justify your answers.
- Calculate $\mathbb{P}[X(1) = 1, X(2) = 0, X(3) = 0, X(4) = 1 \mid X(0) = 0]$.
- Calculate the invariant distribution. Do all initial distributions converge to this invariant distribution? Justify your answer.

Solution:

- The Markov chain is irreducible because $a, b \in (0, 1)$. Also, $P(0, 0) > 0$, so that

$$\gcd\{n > 0 \mid P^n(0, 0) > 0\} = \gcd\{1, 2, 3, \dots\} = 1,$$

which shows that the Markov chain is aperiodic.

We can also notice from the definition of aperiodicity that if a Markov chain has a self loop with nonzero probability, it is aperiodic. In particular, a self loop implies that the smallest number of steps we need to take to get from a state back to itself is 1. In this case, since $P(0, 0) > 0$, we have a self loop with nonzero probability, which makes the Markov chain aperiodic.

- As a result of the Markov property, we know our state at timestep n depends only on timestep $n - 1$. Looking at the transition probabilities, we see that the final expression is

$$P(0, 1) \times P(1, 0) \times P(0, 0) \times P(0, 1) = a(1 - b)(1 - a)a.$$

- The balance equations are

$$\begin{aligned} \begin{cases} \pi(0) = (1 - a)\pi(0) + (1 - b)\pi(1) \\ \pi(1) = a\pi(0) + \pi(2) \end{cases} &\implies \begin{cases} a\pi(0) = (1 - b)\pi(1) \\ \pi(1) = a\pi(0) + \pi(2) \end{cases} \\ &\implies \begin{cases} a\pi(0) = (1 - b)\pi(1) \\ \pi(1) = a\left(\frac{1-b}{a}\pi(1)\right) + \pi(2) \end{cases} \\ &\implies \begin{cases} a\pi(0) = (1 - b)\pi(1) \\ b\pi(1) = \pi(2) \end{cases} \end{aligned}$$

As a side note, these last equations express the equality of the probability of a jump from i to $i + 1$ and from $i + 1$ to i , for $i = 0$ and $i = 1$, respectively. These relations are also called the “detailed balance equations”.

From these equations we find successively that

$$\pi(1) = \frac{a}{1-b}\pi(0) \qquad \pi(2) = b\pi(1) = \frac{ab}{1-b}\pi(0).$$

The normalization equation is

$$\begin{aligned} 1 &= \pi(0) + \pi(1) + \pi(2) = \pi(0) \left(1 + \frac{a}{1-b} + \frac{ab}{1-b} \right) \\ 1 &= \pi(0) \left(\frac{1-b+a+ab}{1-b} \right) \end{aligned}$$

so that

$$\pi(0) = \frac{1-b}{1-b+a+ab}.$$

Thus,

$$\pi(0) = \frac{1-b}{1-b+a+ab} \qquad \pi(1) = \frac{a}{1-b+a+ab} \qquad \pi(2) = \frac{ab}{1-b+a+ab}$$

Or in vector form,

$$\pi = \frac{1}{1-b+a+ab} [1-b \quad a \quad ab].$$

Since the Markov chain is irreducible and aperiodic, all initial distributions converge to this invariant distribution by the fundamental theorem of Markov chains.

5 A Bit of Everything

Note 22

Suppose that X_0, X_1, \dots is a Markov chain with finite state space $S = \{1, 2, \dots, n\}$, where $n > 2$, and transition matrix P . Suppose further that

$$\begin{aligned} P(1, i) &= \frac{1}{n} && \text{for all states } i \text{ and} \\ P(j, j-1) &= 1 && \text{for all states } j \neq 1, \end{aligned}$$

with $P(i, j) = 0$ everywhere else.

- Prove that this Markov chain is irreducible and aperiodic.
- Suppose you start at state 1. What is the distribution of T , where T is the number of transitions until you leave state 1 for the first time?
- Again starting from state 1, what is the expected number of transitions until you reach state n for the first time?
- Again starting from state 1, what is the probability you reach state n before you reach state 2?
- Compute the stationary distribution of this Markov chain.

Solution:

- (a) For any two states i and j , we can consider the path $(i, i-1, \dots, 2, 1, j)$, which has nonzero probability of occurring. Thus, this chain is irreducible. To see that it is aperiodic, observe that $d(1) = 1$, as we have self-loop from state 1 to itself.
- (b) At any given transition, we leave state 1 with probability $\frac{n-1}{n}$, independently of any previous transition. Thus, the distribution is Geometric, with parameter $\frac{n-1}{n}$.
- (c) Suppose that $\beta(i)$ is the expected number of transitions necessary to reach state n for the first time, starting from state i . We have the following first step equations:

$$\begin{aligned}\beta(1) &= 1 + \sum_{j=1}^n \frac{1}{n} \beta(j), \\ \beta(i) &= 1 + \beta(i-1) \quad \text{for } 1 < i < n, \text{ and} \\ \beta(n) &= 0.\end{aligned}$$

We can simplify the second recurrence to

$$\beta(i) = i - 1 + \beta(1) \quad \text{for } 1 < i < n.$$

Substituting this simplified recurrence into the first equation, we get that

$$\beta(1) = 1 + \frac{1}{n} \sum_{i=1}^{n-1} (i - 1 + \beta(1)) = 1 + \frac{1}{n} \sum_{i=1}^{n-1} (i - 1) + \frac{1}{n} \sum_{i=1}^{n-1} \beta(1) = 1 + \frac{(n-2)(n-1)}{2n} + \frac{n-1}{n} \beta(1),$$

which we can solve to get that

$$\beta(1) = \boxed{n + \frac{1}{2}(n-1)(n-2)}.$$

- (d) Suppose that $\alpha(i)$ is the probability that we reach state n before we reach state 2, starting from state i . One immediate observation we can make is that from any state i in $\{2, \dots, n-1\}$, we are guaranteed to see state 2 before state n , as we can only take the path $(i, i-1, \dots, 2, 1)$. Hence, $\alpha(i) = 0$ if $i \in \{2, \dots, n-1\}$. Moreover, $\alpha(n) = 1$, so

$$\alpha(1) = \sum_{i=1}^n \frac{1}{n} \alpha(i) = \frac{1}{n} \alpha(1) + \frac{1}{n},$$

$$\text{hence } \alpha(1) = \boxed{\frac{1}{n-1}}.$$

- (e) We have the balance equations

$$\begin{aligned}\pi(i) &= \frac{1}{n} \pi(1) + \pi(i+1) \quad \text{if } i \neq n, \text{ and} \\ \pi(n) &= \frac{1}{n} \pi(1).\end{aligned}$$

We can collapse the first recurrence to

$$\pi(i) = \frac{n-i}{n}\pi(1) + \pi(n) = \frac{n-i+1}{n}\pi(1),$$

so we can express each stationary probability in terms of the stationary probability of state 1. We can finish by using the normalization equation:

$$\pi(1) + \pi(2) + \cdots + \pi(n) = 1 \implies \frac{1}{n}\pi(1) \sum_{i=1}^n n-i+1 = 1.$$

The last sum can be rearranged to be the sum of the integers from 1 up to n , so we get that

$$\pi(1) = \frac{2}{n+1} \implies \pi = \boxed{\frac{2}{n(n+1)} [n \ n-1 \ \cdots \ 1]}.$$

6 Playing Blackjack

Note 22

Suppose you start with \$1, and at each turn, you win \$1 with probability p , or lose \$1 with probability $1-p$. You will continually play games of Blackjack until you either lose all your money, or you have a total of n dollars.

- Formulate this problem as a Markov chain.
- Let $\alpha(i)$ denote the probability that you end the game with n dollars, given that you started with i dollars.

Notice that for $0 < i < n$, we can write $\alpha(i+1) - \alpha(i) = k(\alpha(i) - \alpha(i-1))$. Find k .

- Using part (b), find $\alpha(i)$, where $0 \leq i \leq n$. (You will need to split into two cases: $p = \frac{1}{2}$ or $p \neq \frac{1}{2}$.)

Hint: Try to apply part (b) iteratively, and look at a telescoping sum to write $\alpha(i)$ in terms of $\alpha(1)$. The formula for the sum of a finite geometric series may be helpful when looking at the case where $p \neq \frac{1}{2}$:

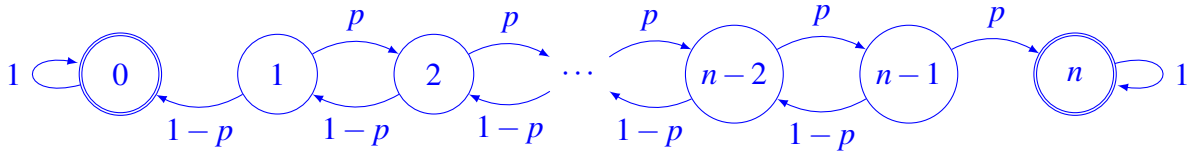
$$\sum_{k=0}^m a^k = \frac{1 - a^{m+1}}{1 - a}.$$

Lastly, it may help to use the value of $\alpha(n)$ to find $\alpha(1)$ for the last few steps of the calculation.

- As $n \rightarrow \infty$, what happens to the probability of ending the game with n dollars, given that you start with i dollars, with the following values of p ?
 - $p > \frac{1}{2}$
 - $p = \frac{1}{2}$
 - $p < \frac{1}{2}$

Solution:

(a) We have the following state transition diagram:



In particular, we have $n + 1$ states, $\{0, 1, 2, \dots, n\}$, where the transition probability from i to $i + 1$ is p , and the transition probability from i to $i - 1$ is $1 - p$. The transition probabilities for $i = 0$ and $i = n$ are edge cases, where we stay in place with probability 1.

(b) If we start with i dollars, this means that we start at state i . The next transition can either be to state $i + 1$ with probability p , or to state $i - 1$ with probability $1 - p$. This means that we have

$$\alpha(i) = p\alpha(i+1) + (1-p)\alpha(i-1).$$

Here, a trick is to expand $\alpha(i) = p\alpha(i) + (1-p)\alpha(i)$. Substituting this in, we can rewrite

$$\begin{aligned} p\alpha(i) + (1-p)\alpha(i) &= p\alpha(i+1) + (1-p)\alpha(i-1) \\ (1-p)(\alpha(i) - \alpha(i-1)) &= p(\alpha(i+1) - \alpha(i)) \\ \alpha(i+1) - \alpha(i) &= \frac{1-p}{p}(\alpha(i) - \alpha(i-1)) \end{aligned}$$

(c) Now that we have a relationship between $\alpha(i+1) - \alpha(i)$ and $\alpha(i) - \alpha(i-1)$, notice that we can iteratively apply the recurrence to get

$$\begin{aligned} \alpha(i+1) - \alpha(i) &= \frac{1-p}{p}(\alpha(i) - \alpha(i-1)) \\ &= \left(\frac{1-p}{p}\right)^2 (\alpha(i-1) - \alpha(i-2)) \\ &\vdots \\ &= \left(\frac{1-p}{p}\right)^i (\alpha(1) - \alpha(0)) \\ &= \left(\frac{1-p}{p}\right)^i \alpha(1) \end{aligned}$$

since $\alpha(0) = 0$ (once we lose all our money, we stop and can never reach n).

Further, notice that we have the telescoping sum

$$[\alpha(i) - \alpha(i-1)] + [\alpha(i-1) - \alpha(i-2)] + \dots + [\alpha(1) - \alpha(0)] = \alpha(i) - \alpha(0) = \alpha(i).$$

This means that we have the summation

$$\begin{aligned}
\alpha(i) &= \sum_{k=0}^{i-1} (\alpha(k+1) - \alpha(k)) \\
&= \sum_{k=0}^{i-1} \left(\frac{1-p}{p} \right)^k \alpha(1) \\
&= \alpha(1) \sum_{k=0}^{i-1} \left(\frac{1-p}{p} \right)^k \\
&= \alpha(1) \cdot \frac{1 - \left(\frac{1-p}{p} \right)^i}{1 - \frac{1-p}{p}}
\end{aligned}$$

[Note that if $p = \frac{1}{2}$, the last step is not valid; in fact, since $\frac{1-p}{p} = 1$, this means that $\alpha(i) = i\alpha(1)$. We'll come back to this case later.]

The previous formula applies for all $0 < i \leq n$, so we can let $i = n$ and simplify to find $\alpha(1)$:

$$\begin{aligned}
1 &= \alpha(n) = \alpha(1) \cdot \frac{1 - \left(\frac{1-p}{p} \right)^n}{1 - \frac{1-p}{p}} \\
\frac{1 - \frac{1-p}{p}}{1 - \left(\frac{1-p}{p} \right)^n} &= \alpha(1)
\end{aligned}$$

Plugging this back in for $\alpha(i)$, we have

$$\alpha(i) = \frac{1 - \frac{1-p}{p}}{1 - \left(\frac{1-p}{p} \right)^n} \cdot \frac{1 - \left(\frac{1-p}{p} \right)^i}{1 - \frac{1-p}{p}} = \frac{1 - \left(\frac{1-p}{p} \right)^i}{1 - \left(\frac{1-p}{p} \right)^n}.$$

Going back to the case where $p = \frac{1}{2}$, we saw that the summation simplifies to $\alpha(i) = i\alpha(1)$. Since $\alpha(n) = 1$, this means that $1 = n\alpha(1)$, or $\alpha(1) = \frac{1}{n}$. This means that we have

$$\alpha(i) = i\alpha(1) = \frac{i}{n}.$$

Together, we have the following formula for any $0 \leq i \leq n$:

$$\alpha(i) = \begin{cases} \frac{1 - \left(\frac{1-p}{p} \right)^i}{1 - \left(\frac{1-p}{p} \right)^n} & p \neq \frac{1}{2} \\ \frac{i}{n} & p = \frac{1}{2} \end{cases}.$$

- (d) (i) If $p > \frac{1}{2}$, then $\frac{1-p}{p} < 1$, and as $n \rightarrow \infty$, the $\left(\frac{1-p}{p} \right)^n$ term in the denominator vanishes. This means that all we're left with is the numerator, and as such

$$\lim_{n \rightarrow \infty} \alpha(i) = 1 - \left(\frac{1-p}{p} \right)^i.$$

(ii) If $p = \frac{1}{2}$, then we know that $\alpha(i) = \frac{i}{n}$. As $n \rightarrow \infty$, this fraction goes to 0, and we have

$$\lim_{n \rightarrow \infty} \alpha(i) = 0.$$

(iii) If $p < \frac{1}{2}$, then $\frac{1-p}{p} > 1$, and as $n \rightarrow \infty$, the $\left(\frac{1-p}{p}\right)^n$ term in the denominator blows up. This means that the denominator tends to $-\infty$, while the numerator remains bounded for any fixed i . This means that the entire fraction tends to 0, i.e.,

$$\lim_{n \rightarrow \infty} \alpha(i) = 0.$$

Note that this problem shows that, even in the case of a fair game (i.e., $p = \frac{1}{2}$), the probability that a gambler wins $\$n$ before going broke tends to zero as $n \rightarrow \infty$. This is one version of the so-called “Gambler’s Ruin” problem. Only in the case where $p > \frac{1}{2}$, i.e., when the game is strictly in the gambler’s favor, does the gambler come out on top with positive probability.