Suponha que desejamos estudar como o comportamento muda sob certas condições experimentais. Pensamos, em particular, em uma sequência de eventos começando com a percepção de um estímulo, seguida pela execução de uma resposta (pressionar uma barra, correr em um labirinto, etc.), e terminando com a ocorrência de um evento ambiental (apresentação de comida, choque elétrico, etc.):

O comportamento é medido pela probabilidade p de que a resposta ocorra durante um determinado intervalo de tempo após o início da sequência. A ideia geral é que p denote o nível de desempenho do sujeito e seja aumentada ou diminuída após cada ocorrência da resposta, conforme os fatores ambientais sejam reforçadores ou inibidores.

Se imaginarmos um experimento no qual um sujeito é repetidamente exposto a essa sequência de eventos (estímulo–resposta–evento ambiental), podemos dividir o experimento em estágios, cada estágio sendo uma tentativa durante a qual o sujeito percorre a sequência. O nível de desempenho do sujeito é então uma função do número de tentativas, denotado por n, e seja p a probabilidade da resposta (durante o intervalo de tempo especificado após o estímulo) na n-ésima tentativa.

O número p_0 será tomado como o valor inicial que descreve a disposição do sujeito em relação à resposta quando ele é introduzido no experimento propriamente dito. A função p é então definida no domínio do conjunto de valores $n = 0, 1, 2, \ldots$. Ao chamar p de uma probabilidade, impomos a normalização:

$$0 \leq p \leq 1, \quad n = 0, 1, 2, \ldots \tag{2.95}$$

o que apenas identifica os extremos de nenhuma resposta e resposta certa com os valores 0 e 1, respectivamente.

Assumimos inicialmente que p_{n+1} depende apenas de p_n e não dos valores anteriores da função p. Em outras palavras, o desempenho do sujeito na tentativa $n+1$, embora dependente do nível de comportamento na tentativa anterior (medido por p_n), é independente do histórico completo até a tentativa n. Isso é conhecido como a propriedade de Markov do modelo.

Seguindo Bush e Mosteller, fazemos a suposição simplificadora de que essa dependência de p_{n+1} em relação a p_n é linear, ou seja, uma linha reta resulta quando p_{n+1} é representada graficamente como função de p_n. A forma de inclinação e intercepto dessa equação é:

$$p_{n+1} = a + mp_n, \quad n = 0, 1, 2, \ldots \tag{2.96}$$

onde a é o intercepto (isto é, o valor de p_{n+1} quando $p_n = 0$) e m é a inclinação da linha (isto é, a variação em p_{n+1} por unidade de variação de p_n).

Para nossos propósitos, é mais conveniente escrever essa relação linear na forma de “ganho-perda”. Introduzimos o parâmetro b pela equação definidora:

$$m = 1 - a - b \tag{2.97}$$

Assim, a relação 2.96 pode ser reescrita como:

$$p_{n+1} = p_n + a(1 - p_n) - bp_n, \quad n = 0, 1, 2, \ldots \tag{2.98}$$

Nota: se conhecemos os valores de a e m, então b é unicamente determinado pela equação 2.97; se a e b são conhecidos, também podemos determinar m.

Assim, as equações (2.96) e (2.98) são formas alternativas, mas equivalentes, da relação linear entre p e p_{n+1}.

Se o nível de desempenho do sujeito na tentativa de número n é dado por p, então $1 - p$ é o aumento máximo possível de desempenho e $-p$ é a diminuição máxima possível ao passar para a tentativa $n + 1$. Isso decorre do fato de que 1 e 0 são os maiores e menores valores possíveis de probabilidade.

A equação (2.98) pode ser interpretada dizendo que a mudança no nível de desempenho, $\Delta p = p_{n+1} - p_n$, é proporcional ao ganho máximo possível e à perda máxima possível. Por esse motivo, a equação (2.98) é chamada de forma "ganho-perda". As constantes de proporcionalidade são a e b, e podemos, portanto, medir com o parâmetro a os eventos ambientais que são reforçadores (por exemplo, apresentação de uma recompensa) e com b os eventos que são inibidores (por exemplo, punição ao sujeito).

As restrições sobre a e b são impostas apenas para garantir que, não importa qual seja o valor de p, consistente com (2.95), o valor seguinte p_{n+1} também estará entre 0 e 1 inclusive.

Se $p = 0$, então $p_{n+1} = a$, de modo que exigimos:

$$0 \leq a \leq 1. \tag{2.99}$$

Se $p = 1$, então $p_{n+1} = 1 - b$, e, portanto, requeremos:

$$0 \leq b \leq 1. \tag{2.100}$$

Mostramos que as condições (2.99) e (2.100) são necessárias para que p_{n+1} esteja entre 0 e 1. Não é difícil mostrar que essas condições também são suficientes (ver Problema 1 da Seção 2.10). Essas são as únicas restrições impostas sobre os parâmetros a e b na equação fundamental de diferenças (2.98).

Assim, $a = 0$ descreve uma situação na qual nenhuma recompensa é dada após a resposta ocorrer, $b = 0$ descreve uma tentativa sem punição, e $a = b$ implica que as medidas de recompensa e punição são iguais.

Citamos Bush e Mosteller: "podemos agora descrever a mudança progressiva na probabilidade de uma resposta em um experimento como o de Graham-Gagné (pista de corrida) ou a caixa de Skinner, nos quais os mesmos eventos ambientais seguem cada ocorrência da resposta."

Vamos considerar um exemplo específico: se $a = 0{,}4$ e $b = 0{,}1$, então a equação (2.98) torna-se:

$$\begin{aligned} p_{n+1} &= p_n + 0{,}4(1 - p_n) - 0{,}1p_n \\ &= 0{,}5p_n + 0{,}4 \end{aligned} \tag{2.101}$$

Se assumirmos $p_0 = 0{,}2$, podemos calcular sucessivamente:

$$p_1 = 0{,}5(0{,}2) + 0{,}4 = 0{,}5,$$
$$p_2 = 0{,}5(0{,}5) + 0{,}4 = 0{,}65.$$

Para resolver a equação (2.101) e obter p_n para todo n, usamos o Teorema 2.7 com $A = 0{,}5$, $B = 0{,}4$. O limite p é dado por:

$$p = \frac{B}{1-A} = \frac{0{,}4}{1-0{,}5} = 0{,}8. \tag{1}$$

A solução geral é então:

$$p_n = (0{,}5)^n(p_0 - 0{,}8) + 0{,}8. \tag{2.102}$$

Essa solução mostra exatamente como o nível de desempenho varia com o número de tentativas. Como $0 < A < 1$ e $p_0 < p$, sabemos (ver Tabela 2.2) que a sequência $\{p_n\}$ é monótona crescente com limite $p = 0{,}8$.

Assim, para tentativas repetidas em que recompensa e punição têm peso na razão 4:1 (como em $a = 0{,}4$ e $b = 0{,}1$), as probabilidades de resposta e de não resposta, p e $1-p$, se aproximam de valores limites cuja razão é a mesma.

Voltando ao caso geral, notamos que (2.98) pode ser reescrita na forma padrão:

$$p_{n+1} = (1-a-b)p_n + a, \quad n = 0, 1, 2, \ldots \tag{2.103}$$

Reconhecemos essa equação como uma equação de diferenças linear de primeira ordem com coeficientes constantes. De fato, usando a notação do Teorema 2.7, temos:

$$A = 1 - a - b, \quad B = a,$$

e, portanto, o valor limite p é:

$$p = \frac{B}{1-A} = \frac{a}{a+b}, \quad \text{se } a+b \neq 0. \tag{2.104}$$

Temos então a solução geral:

$$p_n = \begin{cases} (1-a-b)^n(p_0 - p) + p, & \text{se } a+b \neq 0, \\ p_0, & \text{se } a+b = 0, \end{cases} \tag{2.105}$$

À luz das condições (2.99) e (2.100), a constante $A = 1-a-b$ está entre -1 e 1, com os extremos atingidos apenas se a e b forem ambos iguais a 0 ou ambos iguais a 1.

Se $a = b = 1$, então $A = -1$ e a sequência $\{p_n\}$ oscila finitamente entre os valores p_0 e $1 - p_0$. Mas em todos os outros casos a sequência $\{p_n\}$ converge, para o limite p_0 se $a = b = 0$, e para o limite p caso contrário.

Se $0 < a+b < 1$, então $0 < A < 1$ e $\{p_n\}$ é monótona:

- decrescente, se $p_0 > p$,

- crescente, se $p_0 < p$,
- constante, se $p_0 = p$.

Se $1 < a + b < 2$, então $-1 < A < 0$ e $\{p_n\}$ é uma sequência oscilatória amortecida com limite p. O caso especial $a+b = 0$ gera uma sequência constante (com valor p_0), e $a + b = 1$ gera uma sequência em que todos os elementos são iguais a p.

Casos Especiais

Concluímos com dois casos especiais:

1. $a = 0$
2. $a = b$

Caso (1): Assume-se que nenhuma recompensa é dada após a ocorrência da resposta. A equação de diferenças (2.98) torna-se:

$$p_{n+1} = (1 - b)p_n \tag{2}$$

com solução:

$$p_n = (1 - b)^n p_0, \quad n = 0, 1, 2, \ldots \tag{3}$$

Esta é uma equação que descreve a diminuição constante da probabilidade de resposta (à medida que $n \to \infty$) a partir da probabilidade inicial p_0. Ao representar p como função de n, obtemos uma curva de extinção experimental (ver discussão no Capítulo 0, Exemplo 1).

Caso (2): Quando $a = b$, as medidas de reforço e punição são iguais. Descartando os casos extremos $a = b = 0$ e $a = b = 1$, temos que a quantidade $(1 - a - b) \to 0$ conforme $n \to \infty$, e a equação (2.105) mostra que o limite de p_n é:

$$p = \frac{a}{a + b} = \frac{a}{2a} = \frac{1}{2}.$$

Ou seja, a resposta tende a ocorrer (no intervalo de tempo especificado após o estímulo) em metade das tentativas. O equilíbrio entre forças de recompensa e punição produz, no longo prazo, uma simetria correspondente no desempenho.

Problemas 2.10

1. Se $p_{n+1} = p_n + a(1 - p_n) - bp_n$ e $0 \leq p_n \leq 1$, mostre que, se $0 \leq a \leq 1$ e $0 \leq b \leq 1$, então também $0 \leq p_{n+1} \leq 1$.
   [**Dica:** Estabeleça as desigualdades:
   $p + a(1 - p) - bp \leq p + 1(1 - p) - 0 = 1$, quando $p = 0$;
   $p + a(1 - p) - bp \geq p + 0(1 - p) - p = 0$, quando $p = 1$.]

2. Usando a equação (2.98) com $a = 0{,}5$ e $b = 0{,}2$, calcule os valores de p_{n+1} para:

$$p_n = 0,\ 0{,}1,\ 0{,}2,\ \ldots,\ 0{,}9,\ 1$$

e trace o gráfico da reta, com p_n no eixo horizontal e p_{n+1} no eixo vertical.

3. (a) Resolva a equação de diferenças (2.98) no caso $p_0 = 0{,}5$, $a = 0{,}1$, e $b = 0{,}4$, e mostre que, quando as medidas de reforço e punição estão na razão 1:4, as probabilidades limites de resposta e não resposta estão na mesma razão.

   (b) Generalize esse resultado mostrando que, quando as medidas de reforço e punição estão na razão $a : b$, as probabilidades limites de resposta e não resposta estão na mesma razão.

4. Com n no eixo horizontal e p_n no eixo vertical, trace os gráficos da função p_n nos seguintes casos:

   (a) $a = 0,\ b = 0{,}2$

   (b) $a = 0,\ b = 0{,}5$

   (c) $a = 0,\ b = 0{,}8$

   Suponha $p_0 = 0{,}5$ e mostre que quanto maior for b, mais rapidamente $p_n \to 0$.

5. Repita o problema anterior com $p_0 = 0{,}5$, mas agora:

   (a) $a = 0{,}2,\ b = 0$

   (b) $a = 0{,}5,\ b = 0$

   (c) $a = 0{,}8,\ b = 0$