2. 解：(1) 状态空间 $S=\{X_n=-2,\ X_n=-1,\ X_n=0,\ X_n=1,\ X_n=2\}$

状态转移矩阵：

$$
\begin{array}{c c}
 & \begin{array}{ccccc} -2 & -1 & 0 & 1 & 2 \end{array} \\
\begin{array}{c} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{array} &
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
q & r & p & 0 & 0 \\
0 & q & r & p & 0 \\
0 & 0 & q & r & p \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\end{array}
$$

(2) 设在甲积1分的情况下，再赛两局可以结束

$S_t=[0\quad 0\quad 0\quad 1\quad 0]^T$

下一局：$S_{t+1}=P^T S_t=[0\quad 0\quad q\quad r\quad p]^T$

经过两局：$S_{t+2}=P^T S_{t+1}=[0\quad q^2\quad 2qr\quad r^2+pq\quad p+pr]^T$

故再赛两局可以结束比赛的概率为 $p+pr-p=pr$
(刚好两局)

3. 解： $G_t = R_{t+1} + \gamma R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+1+k}$

故 $G_0 = \lim_{k\to\infty} R_1 + \gamma R_2 + \gamma^2 R_3 + \dots \gamma^k R_{k+1}$

$= \lim_{k\to\infty} (6 + 6\gamma + 6\gamma^2 + \dots + 6\gamma^k) - 4$

$= \lim_{k\to\infty} \frac{6(\gamma^{k+1}-1)}{\gamma-1} - 4$

$= 6 \times \frac{0-1}{0.9-1} - 4$

$= 56$

故 $G_1 = \lim_{k\to\infty} R_2 + \gamma R_3 + \dots + \gamma^k R_{k+2}$

$= 6 \times \lim_{k\to\infty} (1 + \gamma + \dots + \gamma^k)$

$= \lim_{k\to\infty} \frac{6(\gamma^{k+1}-1)}{\gamma-1}$

$= 60$

4. 解:

① 认为 $R_t=-1$ 对所有行动状态成立，即 $1\to1$ 时 $R_t=-1$ 也成立

$$q_\pi(4,\text{left}) = r_4^{\text{left}} + \sum_{s'\in S} p_{4s'}^{\text{left}} v_\pi(s') = -1 + v_\pi(3)$$

对于状态价值函数：

$v_\pi(1)$、$v_\pi(5)$、$v_\pi(3)$、$v_\pi(7)$ 具有如下相同形式：

$$v_\pi(1) = -\frac{1}{4} + \frac{v_\pi(1)-1}{4} + \frac{v_\pi(2)-1}{4} + \frac{v_\pi(4)-1}{4}$$

$v_\pi(2)$、$v_\pi(6)$ 具有如下相同形式：

$$v_\pi(2) = \frac{v_\pi(2)-1}{2} + \frac{v_\pi(3)-1}{4} + \frac{v_\pi(1)-1}{4}$$

$v_\pi(4)$ 具有如下形式：

$$v_\pi(4) = \frac{v_\pi(1)-1}{4} + \frac{v_\pi(5)-1}{4} + \frac{v_\pi(3)-1}{4} + \frac{v_\pi(7)-1}{4}$$

得：$v_\pi(1) = v_\pi(5) = v_\pi(3) = v_\pi(1) = -7$

$v_\pi(2) = v_\pi(6) = -9$

$v_\pi(4) = -8$

故 $q_\pi(4,\text{left}) = -1 + (-7) = -8$

$q_\pi(7,\text{right}) = r_7^{\text{right}} = -1$

② 认为 $R_t=-1$ 只在状态发生转移时成立，做动作但状态不变认为 $R_t=0$ ex: 从1到1 R_t 记为0)

$$q_\pi(4,\text{left}) = r_4^{\text{left}} + \sum_{s\in S} P_{4s'}^{\text{left}} V_\pi(s') = -1 + V_\pi(3)$$

对于状态价值函数：

$V_\pi(1)$、$V_\pi(5)$、$V_\pi(3)$、$V_\pi(7)$ 具有如下相同形式：

$$V_\pi(1) = -\frac{1}{4} + \frac{V_\pi(1)}{4} + \frac{V_\pi(2)-1}{4} + \frac{V_\pi(5)-1}{4}$$

$V_\pi(2)$、$V_\pi(6)$ 具有如下相同形式：

$$V_\pi(2) = \frac{V_\pi(2)}{2} + \frac{V_\pi(3)-1}{4} + \frac{V_\pi(1)-1}{4}$$

$V_\pi(4)$ 具有如下形式：

$$V_\pi(4) = \frac{V_\pi(1)-1}{4} + \frac{V_\pi(5)-1}{4} + \frac{V_\pi(3)-1}{4} + \frac{V_\pi(7)-1}{4}$$

得：$V_\pi(1) = V_\pi(5) = V_\pi(3) = V_\pi(1) = -5$

$V_\pi(2) = V_\pi(4) = V_\pi(6) = -6$

故 $q_\pi(4,\text{left}) = -1 + (-5) = -6$

$q_\pi(7,\text{right}) = r_7^{\text{right}} = -1$