

OPTIMAL STOPPING IN A STOCHASTIC GAME

BAHAR KAYNAR

*Department of Econometrics and Operations Research
VU University Amsterdam
Amsterdam, The Netherlands
E-mail: bkaynar@feweb.vu.nl*

In this article we consider a stochastic game in which each player draws one or two random numbers between 0 and 1. Players can decide to stop after the first draw or to continue for a second draw. The decision is made without knowing the other players' numbers or whether the other players continue for a second draw. The object of the game is to have the highest total score without going over 1. In the article, we will characterize the optimal stopping rule for each player.

1. INTRODUCTION

In the popular TV game show “The Price is Right,” the Showcase Showdown game of chance is played on every show. In the game each of three players in turn spins the wheel once or twice attaining some total score and then waits for the results of the succeeding players' spins. The object of the game is to have the highest score, from one or two spins, without going over a given upper limit. This game of chance has been analyzed in Coe and Butterworth [1] and Tijms [2].

This article considers a variant of this stochastic game, in which each player has no information about the results and actions of the other players. Each player chooses one or two random numbers between 0 and 1. The player can decide to stop after the first draw or to continue for a second draw. The decision must be made without knowing what the other players have done. The object of the game is to have the highest total score without going over 1. In case the total scores of all players exceed 1, the winner is the player whose score is closest to 1. What stopping rule should a player use in order to maximize its probability of winning?

It is obvious that each player uses a rule characterized by a single threshold value v : continue for a second draw if the first draw gives a number less than v , otherwise

stop after the first draw. In Section 2, we first consider the case of two players and we give the optimal stopping level that ensures a player winning at least 50% of the time whatever the other player is doing. The solution for the case of three players will be derived in Section 3. Finally, we discuss the general case of many players in Section 4.

2. TWO-PERSON GAME

For the case of two players, let $P(a, b)$ denote the winning probability of Player A when Player A uses a threshold a and Player B uses a threshold b . Player A wants to use the threshold $a = a_0$, where a_0 attains the maximum in $\max_a \min_b P(a, b)$. It suffices to determine $P(a, b)$ for $a \geq b$. By a symmetry argument, we have

$$P(a, b) = 1 - P(b, a) \quad \text{for } b \geq a.$$

For fixed values of the thresholds a and b of the two players, define the random variable \mathbf{X}_A as the total score of Player A. Also let $F_A(x) = P(\mathbf{X}_A \leq x)$ for $0 < x < 1 + a$, and let $f_A(x)$ be the probability density of $F_A(x)$. Similarly, $\mathbf{X}_B, F_B(x)$, and $f_B(x)$ are defined. By a simple conditioning argument,

$$P(\mathbf{X}_A \leq x) = \int_0^x (x - y) dy = \frac{1}{2}x^2 \quad \text{for } 0 < x \leq a.$$

In addition, it is readily seen by conditioning that

$$P(\mathbf{X}_A > x) = \begin{cases} 1 - x + \int_0^a (1 - (x - y)) dy & \text{for } a < x \leq 1 \\ \int_{x-1}^a (1 - (x - y)) dy & \text{for } 1 < x < 1 + a. \end{cases}$$

This leads to

$$F_A(x) = P(\mathbf{X}_A \leq x) = \begin{cases} \frac{1}{2}x^2 & \text{for } 0 < x \leq a \\ (1 + a)x - a - \frac{1}{2}a^2 & \text{for } a < x \leq 1 \\ \frac{1}{2} - \frac{1}{2}(x - a)^2 + x - a & \text{for } 1 < x < 1 + a. \end{cases} \tag{2.1}$$

By differentiating, we get

$$f_A(x) = \begin{cases} x & \text{for } 0 < x \leq a \\ 1 + a & \text{for } a < x \leq 1 \\ 1 + a - x & \text{for } 1 < x < 1 + a. \end{cases} \tag{2.2}$$

The formulas for $F_B(x) = P(\mathbf{X}_B \leq x)$ and $f_B(x)$ follow by replacing a by b in (2.1) and (2.2). By conditioning on the value of \mathbf{X}_A , for $a \geq b$,

$$\begin{aligned} P(a, b) &= \int_0^a P(A \text{ beats } B \mid \mathbf{X}_A = x) x \, dx \\ &\quad + \int_a^1 P(A \text{ beats } B \mid \mathbf{X}_A = x) (1 + a) \, dx \\ &\quad + \int_1^{1+a} P(A \text{ beats } B \mid \mathbf{X}_A = x) (1 + a - x) \, dx. \end{aligned}$$

This can be weighted out as

$$\begin{aligned} P(a, b) &= \int_0^b [P(\mathbf{X}_B \leq x) + P(\mathbf{X}_B > 1)] x \, dx \\ &\quad + \int_b^a [P(\mathbf{X}_B \leq x) + P(\mathbf{X}_B > 1)] x \, dx \\ &\quad + \int_a^1 [P(\mathbf{X}_B \leq x) + P(\mathbf{X}_B > 1)] (1 + a) \, dx \\ &\quad + \int_1^{1+b} P(1 < x \leq \mathbf{X}_B) (1 + a - x) \, dx \\ &\quad + \int_{1+b}^{1+a} P(1 < x \leq \mathbf{X}_B) (1 + a - x) \, dx. \end{aligned}$$

After some algebra, we were able to represent $P(a, b)$ in the following form:

$$P(a, b) = \frac{1}{2} - \frac{1}{6} (a - b) (a^2 b + a^2 + ab^2 + b^2 + ab + 3a - 3) \quad \text{for } a \geq b.$$

By symmetry argument, we have $P(a, b) = 1 - P(b, a)$ for $a \leq b$, which gives

$$P(a, b) = \frac{1}{2} - \frac{1}{6} (a - b) (a^2 b + a^2 + ab^2 + b^2 + ab + 3b - 3) \quad \text{for } a \leq b.$$

This representation is crucial for our analysis. To find a_0 , we reason as follows. For the threshold a_0 we should have $P(a_0, b) \geq \frac{1}{2}$ for all b . This gives

$$\begin{aligned} a_0^2 b + a_0^2 + a_0 b^2 + b^2 + a_0 b + 3a_0 - 3 &\leq 0 \quad \text{for } a_0 \geq b, \\ a_0^2 b + a_0^2 + a_0 b^2 + b^2 + a_0 b + 3b - 3 &\geq 0 \quad \text{for } a_0 \leq b. \end{aligned} \tag{2.3}$$

If Player B chooses $b = a_0$ as the threshold value, we obtain from (2.3) that

$$2a_0^3 + 3a_0^2 + 3a_0 - 3 = 0. \tag{2.4}$$

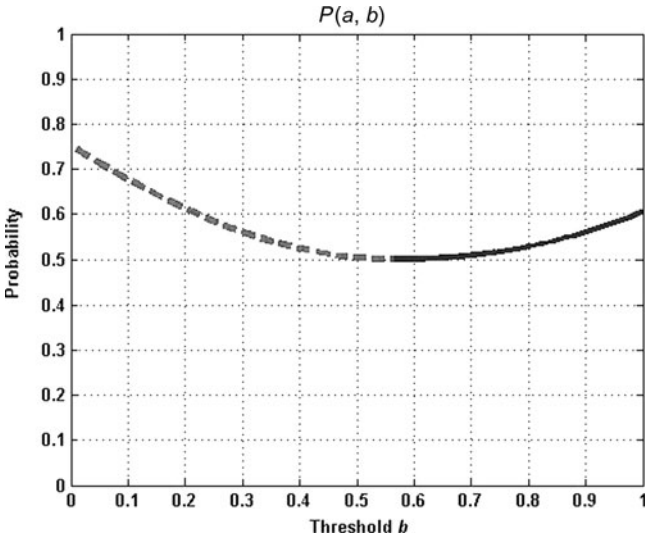


FIGURE 1. $P(a_0, b)$.

By solving this equality, we get the optimal threshold value as

$$a_0 = \frac{1}{2} \left\{ -1 - \frac{1}{\sqrt[3]{8 + \sqrt{65}}} + \sqrt[3]{8 + \sqrt{65}} \right\} = 0.563385.$$

It can be observed from Figure 1 that $P(a_0, b)$ is always greater than .5 for all values of b with $b \neq a_0$.

3. EXTENSION TO THREE PLAYERS

Let us try to extend the game to three people and see how the strategy changes as the number of players increases. Of course, the question should be updated to what threshold value should Player A choose to win with probability more than 1/3 all the time. The random variables \mathbf{X}_A , \mathbf{X}_B , and \mathbf{X}_C denote the total scores of the players. Also, let $P(a, b, c)$ denote the winning probability of Player A when the players have threshold values a , b , and c , respectively. With the same type of approach as used in the previous section, we calculate the winning probability of A, by conditioning on the value of \mathbf{X}_A :

$$\begin{aligned} P(a, b, c) &= P(A \text{ wins}) \\ &= \int_0^{1+a} P(A \text{ wins} \mid \mathbf{X}_A = x) f_A(x) dx. \end{aligned}$$

Let us first consider the case $a \geq b \geq c$:

$$\begin{aligned} P(a, b, c) &= \int_0^a P(A \text{ wins} \mid \mathbf{X}_A = x)x \, dx \\ &\quad + \int_a^1 P(A \text{ wins} \mid \mathbf{X}_A = x)(1 + a) \, dx \\ &\quad + \int_1^{1+a} P(A \text{ wins} \mid \mathbf{X}_A = x)(1 + a - x) \, dx. \end{aligned}$$

In order for Player A to win the game with a total score x with $x \leq 1$, the total score of Player B should be either less than or equal to x or greater than 1. This should also hold for Player C . Hence, for $0 < x \leq 1$,

$$\begin{aligned} P(A \text{ wins} \mid \mathbf{X}_A = x) &= P(\mathbf{X}_B \leq x)P(\mathbf{X}_C \leq x) + P(\mathbf{X}_B \leq x)P(\mathbf{X}_C > 1) \\ &\quad + P(\mathbf{X}_B > 1)P(\mathbf{X}_C \leq x) + P(\mathbf{X}_B > 1)P(\mathbf{X}_C > 1) \end{aligned}$$

using the independency of \mathbf{X}_B and \mathbf{X}_C . Also, for $1 < x < 1 + a$,

$$P(A \text{ wins} \mid X_A = x) = P(\mathbf{X}_B \geq x)P(\mathbf{X}_C \geq x).$$

For the case $a \geq b \geq c$, the winning probability of A can be obtained as

$$\begin{aligned} P(a, b, c) &= \int_0^c P(A \text{ wins} \mid X_A = x)x \, dx \\ &\quad + \int_c^b P(A \text{ wins} \mid X_A = x)x \, dx \\ &\quad + \int_b^a P(A \text{ wins} \mid X_A = x)x \, dx \\ &\quad + \int_a^1 P(A \text{ wins} \mid X_A = x)(1 + a) \, dx \\ &\quad + \int_1^{1+a} P(A \text{ wins} \mid X_A = x)(1 + a - x) \, dx. \end{aligned}$$

After tedious algebra, we find for $P(a, b, c)$ with $a \geq b \geq c$ the expression

$$P(a, b, c) = \frac{1}{3} + \frac{1}{120}(a - b)f_1(a, b, c) + \frac{1}{120}(b - c)g_1(a, b, c)$$

for some functions f_1 and g_1 . Working out the other cases in the same way, we obtain

$$\begin{aligned}
 a \geq b \geq c: \quad & P(a, b, c) = \frac{1}{3} + \frac{1}{120}(a - b)f_1(a, b, c) + \frac{1}{120}(b - c)g_1(a, b, c), \\
 a \geq c \geq b: \quad & P(a, b, c) = \frac{1}{3} + \frac{1}{120}(a - b)f_2(a, b, c) + \frac{1}{120}(b - c)g_2(a, b, c), \\
 b \geq a \geq c: \quad & P(a, b, c) = \frac{1}{3} + \frac{1}{120}(a - b)f_3(a, b, c) + \frac{1}{120}(b - c)g_3(a, b, c), \\
 c \geq a \geq b: \quad & P(a, b, c) = \frac{1}{3} + \frac{1}{120}(a - b)f_4(a, b, c) + \frac{1}{120}(b - c)g_4(a, b, c), \\
 b \geq c \geq a: \quad & P(a, b, c) = \frac{1}{3} + \frac{1}{120}(a - b)f_5(a, b, c) + \frac{1}{120}(b - c)g_5(a, b, c), \\
 c \geq b \geq a: \quad & P(a, b, c) = \frac{1}{3} + \frac{1}{120}(a - b)f_6(a, b, c) + \frac{1}{120}(b - c)g_6(a, b, c).
 \end{aligned}$$

The explicit expressions for the functions f_1, \dots, f_6 and g_1, \dots, g_6 are given in the Appendix. As an accuracy check, the complex formulas are confirmed by simulation by taking several choices for a, b , and c . Using a similar approach as in the previous section, we will find the reduced polynomial for a_0 . Player A wants to win the game with a probability of at least $1/3$ for all values of b and c . In other words, for the optimal threshold a_0 , we have $P(a_0, b, c) \geq \frac{1}{3}$ for all b and c . This gives

$$\frac{1}{120}(a_0 - b)f_i(a_0, b, c) + \frac{1}{120}(b - c)g_i(a_0, b, c) \geq 0$$

for $i = 1, 2, \dots, 6$. If Players B and C play with the same threshold value $b = c$, the first two and the last two cases for $P(a, b, c)$ boil down to

$$\frac{1}{120}(a_0 - b)f_i(a_0, b, b) \geq 0, \quad i = 1, 2, 5, 6.$$

This implies

$$\begin{aligned}
 f_1(a_0, b, b) \geq 0 \quad \text{and} \quad f_2(a_0, b, b) \geq 0 \quad \text{for } a_0 \geq b, \\
 f_5(a_0, b, b) \leq 0 \quad \text{and} \quad f_6(a_0, b, b) \leq 0 \quad \text{for } a_0 \leq b.
 \end{aligned} \tag{3.1}$$

Taking $b = a_0$ in (3.1) and using the explicit expressions for $f_i(a, b, c)$ in the Appendix, we find

$$f_i(a_0, a_0, a_0) = 2(20 - 20a_0 + 20a_0^2 - 20a_0^3 - 40a_0^4 - 17a_0^5) \quad \text{for } i = 1, 2, 5, 6.$$

These expressions together with (3.1) give us

$$20 - 20a_0 + 20a_0^2 - 20a_0^3 - 40a_0^4 - 17a_0^5 = 0. \tag{3.2}$$

By solving (3.2), for the case of three players we obtain the optimal threshold value

$$a_0 = 0.660527.$$

It is interesting to note that we can also obtain (3.2) by using the third and the fourth cases for $P(a, b, c)$. If Player C selects $c = a_0$ as the threshold value, we obtain

$$\begin{aligned} \frac{1}{120}(a_0 - b)(f_3(a_0, b, a_0) - g_3(a_0, b, a_0)) &\geq 0 \quad \text{for } b \geq a_0, \\ \frac{1}{120}(a_0 - b)(f_4(a_0, b, a_0) - g_4(a_0, b, a_0)) &\geq 0 \quad \text{for } b \leq a_0 \end{aligned}$$

or, equivalently,

$$\begin{aligned} f_3(a_0, b, a_0) - g_3(a_0, b, a_0) &\leq 0 \quad \text{for } b \geq a_0, \\ f_4(a_0, b, a_0) - g_4(a_0, b, a_0) &\geq 0 \quad \text{for } b \leq a_0. \end{aligned}$$

When Player B also chooses $b = a_0$, the two expressions lead to

$$\begin{aligned} f_i(a_0, a_0, a_0) - g_i(a_0, a_0, a_0) &= 20 - 20a_0 + 20a_0^2 - 20a_0^3 - 40a_0^4 - 17a_0^5 \\ &= 0 \quad \text{for } i = 3, 4. \end{aligned}$$

4. WHAT ABOUT MORE THAN THREE PLAYERS?

In order to obtain $P(A \text{ wins})$ for the case of n players, we have to distinguish between $n!$ possible orderings of the threshold values d_1, \dots, d_n of the players. For the i th possible ordering, the following general form of $P(A \text{ wins})$ holds

$$\begin{aligned} P(A \text{ wins}) &= \frac{1}{n} + C_1(d_1 - d_2)f_i^1(d_1, \dots, d_n) \\ &\quad + \dots + C_{n-1}(d_{n-1} - d_n)f_i^{n-1}(d_1, \dots, d_n), \quad i = 1, \dots, n!. \end{aligned}$$

Using Mathematica, we could obtain explicit expression for the functions $f_i^j(d_1, \dots, d_n)$ up to $n = 10$. The next step is to compute the optimal threshold value a_0 from these expressions. To do this, we used a shortcut in our analysis. The following remarkable observation can be made for the case of three players. If we plug $a = b = c = a_0$ into the expressions for the functions $f_i(d_1, \dots, d_n)$ and $g_i(d_1, \dots, d_n)$ in the Appendix, we get, for all i , the following reduced polynomial for a_0 :

$$\begin{aligned} f_i(a_0, a_0, a_0) &= \frac{1}{120}(40 - 40a_0 + 40a_0^2 - 40a_0^3 - 80a_0^4 - 34a_0^5), \\ g_i(a_0, a_0, a_0) &= \frac{1}{120}(20 - 20a_0 + 20a_0^2 - 20a_0^3 - 40a_0^4 - 17a_0^5). \end{aligned}$$

The same observation applies to the case of $n = 2$ players. This observation was used to get the polynomial for a_0 in the general case of $n > 3$. The power of Mathematica enabled us to obtain the reduced polynomial up to $n = 10$ players. The polynomial for the case of n players is of degree $2n - 1$. In Table 1 we only give the reduced

TABLE 1. Reduced Polynomials for n Players

n	Reduced polynomial
2	$3 - 3a_0 - 3a_0^2 - 2a_0^3$
3	$20 - 20a_0 + 20a_0^2 - 20a_0^3 - 40a_0^4 - 17a_0^5$
4	$14 - 14a_0 + 14a_0^2 - 14a_0^3 + 14a_0^4 - 14a_0^5 - 42a_0^6 - 13a_0^7$
5	$144 - 144a_0 + 144a_0^2 - 144a_0^3 + 144a_0^4 - 144a_0^5 + 144a_0^6 - 144a_0^7 - 576a_0^8 - 139a_0^9$

TABLE 2. Optimal Threshold Values for n Players

n	a_0	a^*
2	0.563385	0.53209
3	0.660527	0.64865
4	0.717991	0.71145
5	0.756631	0.75225
6	0.784680	0.78141
7	0.806119	0.80353
8	0.823129	0.820995
9	0.837011	0.835209
10	0.848594	0.847044

polynomial for $n = 2, 3, 4, 5$, but we have computed the optimal threshold a_0 for up to $n = 10$ players. The shortcut we made to obtain a_0 was empirically validated by simulation experiments.

The optimal threshold value a_0 is in Table 2 for $n = 2, 3, \dots, 10$. It is interesting to compare the optimal threshold values with the optimal stopping level that is used by the first player in the stochastic game in which each of the players in turn chooses one or two random numbers between 0 and 1 and then waits for the results of the succeeding players' draws. In this continuous version of the Showcase Showdown game, we have for the case of n players that the optimal stopping level, a^* , of the first player is the solution of

$$a^{2n-2} = \frac{1}{2n-1} (1 - a^{2n-1})$$

on $(0, 1)$, see Tijms [2]. The numerical values of a^* are given in the last column of Table 2. In comparison, as the number of players increases, the optimal values get closer to each other.

Acknowledgment

I am indebted to professor Henk Tijms for suggesting this problem.

References

1. Coe, P.R. & Butterworth, W. (1995). Optimal stopping in "The Showcase Showdown," *The American Statistician* 49: 271–275.
2. Tijms, H.C. (2007). *Understanding probability*, 2nd ed. Cambridge: Cambridge University Press.

APPENDIX

For the case of three players, the exact formulas for f_i and g_i as functions of threshold values a , b , and c are as follows:

$$f_1(a, b, c) = \frac{1}{120} (40 - 40a^2 - 10a^3 - 20b + 20ab - 30a^2b - 10a^3b + 20b^2 - 30ab^2 - 10a^2b^2 - 30b^3 - 10ab^3 - 10b^4 - 20c + 60ac - 20a^2c - 10a^3c - 20bc + 40abc - 10a^2bc - 10a^3bc + 40b^2c - 10ab^2c - 10a^2b^2c - 10b^3c - 10ab^3c - 10b^4c + 10b^2c^3 - 5bc^4 + c^5),$$

$$g_1(a, b, c) = \frac{1}{120} (20 - 20b + 20b^2 - 20b^3 - 8b^4 - 13b^3c - 8b^4c - 13b^2c^2 - 8b^3c^2 - 3bc^3 + 2b^2c^3 - 3c^4 - 3bc^4),$$

$$f_2(a, b, c) = \frac{1}{120} (40 - 40a^2 - 10a^3 - 20b + 20ab - 30a^2b - 10a^3b + 20b^2 - 30ab^2 - 10a^2b^2 - 30b^3 - 10ab^3 - 10b^4 + b^5 - 20c + 60ac - 20a^2c - 10a^3c - 20bc + 40abc - 10a^2bc - 10a^3bc + 40b^2c - 10ab^2c - 10a^2b^2c - 10b^3c - 10ab^3c - 15b^4c + 10b^3c^2),$$

$$g_2(a, b, c) = \frac{1}{120} (20 - 20b + 20b^2 - 30b^3 - 7b^4 + b^5 + 10b^2c - 17b^3c - 12b^4c + 10bc^2 - 7b^2c^2 - 2b^3c^2 - 10c^3 - 7bc^3 - 2b^2c^3 - 2c^4 - 2bc^4),$$

$$f_3(a, b, c) = \frac{1}{120} (40 - 15a^3 - 3a^4 - 20b - 15a^2b - 3a^3b - 45ab^2 - 13a^2b^2 - 25b^3 - 13ab^3 - 8b^4 - 20c + 20a^2c - 10a^3c - 3a^4c + 40bc + 20abc - 10a^2bc - 3a^3bc + 20b^2c - 10ab^2c - 13a^2b^2c - 10b^3c - 13ab^3c - 8b^4c + 10b^2c^3 - 5bc^4 + c^5),$$

$$g_3(a, b, c) = \frac{1}{120} (20 - 20b + 20b^2 - 20b^3 - 8b^4 - 13b^3c - 8b^4c - 13b^2c^2 - 8b^3c^2 - 3bc^3 + 2b^2c^3 - 3c^4 - 3bc^4),$$

$$f_4(a, b, c) = \frac{1}{120} (40 - 15a^3 - 3a^4 - 20b + 5a^2b - 13a^3b - 3a^4b + 5ab^2 - 13a^2b^2 - 3a^3b^2 + 5b^3 - 13ab^3 - 3a^2b^3 - 13b^4 - 3ab^4 - 2b^5 - 20c + 40bc - 5b^4c - 30ac^2 - 10a^2c^2 - 30bc^2 - 10abc^2 - 10a^2bc^2 - 10b^2c^2 - 10ab^2c^2 + 20c^3 + 5c^4 + 5bc^4),$$

$$g_4(a, b, c) = \frac{1}{120} (20 - 20b + 5b^3 - 10b^4 - 2b^5 + 40bc + 5b^2c - 10b^3c - 5b^4c - 20c^2 - 25bc^2 - 10b^2c^2 - 5b^3c^2 - 5c^3 - 10bc^3 - 5b^2c^3),$$

$$f_5(a, b, c) = \frac{1}{120} (40 - 6a^4 - 20b - 6a^3b - 3a^4b - 16a^2b^2 - 3a^3b^2 + 20b^3 - 16ab^3 - 3a^2b^3 - 11b^4 - 3ab^4 - 3b^5 - 20c - 3a^4c + 40bc + 7a^3bc - 60b^2c - 3a^2b^2c - 3ab^3c + 2b^4c - 10a^2c^2 - 10abc^2 - 10a^2bc^2 - 10b^2c^2 - 10ab^2c^2 - 10b^3c^2 + 10b^2c^3 + 5c^4 + c^5),$$

$$g_5(a, b, c) = \frac{1}{120} (20 - 20b + 20b^2 + 25b^3 - 11b^4 - 3b^5 - 35b^2c - 6b^3c - b^4c - 5bc^2 - 16b^2c^2 - 11b^3c^2 - 5c^3 - 6bc^3 - b^2c^3 - c^4 - bc^4),$$

$$f_6(a, b, c) = \frac{1}{120} (40 - 6a^4 - 20b - 6a^3b - 3a^4b - 16a^2b^2 - 3a^3b^2 - 16ab^3 - 3a^2b^3 - 11b^4 - 3ab^4 - 2b^5 - 20c - 3a^4c + 40bc + 7a^3bc - 3a^2b^2c - 3ab^3c - 3b^4c - 10a^2c^2 - 60bc^2 - 10abc^2 - 10a^2bc^2 - 10b^2c^2 - 10ab^2c^2 + 20c^3 + 5c^4 + 5bc^4),$$

$$g_6(a, b, c) = \frac{1}{120} (20 - 20b + 5b^3 - 10b^4 - 2b^5 + 40bc + 5b^2c - 10b^3c - 5b^4c - 20c^2 - 25bc^2 - 10b^2c^2 - 5b^3c^2 - 5c^3 - 10bc^3 - 5b^2c^3).$$