

5 Metodologia

Este capítulo será destinado ao método para determinar a apreciação de reservas de óleo e gás incluindo as incertezas das funções de apreciação, determinando um intervalo de confiança representado pelo limites inferior e superior em relação à média ajustando, se possível, a um modelo, determinado uma função de apreciação.

Será abordado o método utilizado para determinar o intervalo de confiança em relação ao valor médio da apreciação das reservas de campos para cada ano após o início da produção e ainda alguns conceitos estatísticos relacionados com intervalo de confiança.

5.1. Associação da Incerteza à Função de Crescimento

A partir da revisão bibliográfica sobre apreciação de reservas no Capítulo 4, verifica-se que existem diversos métodos utilizados para se determinar a apreciação. Entretanto, pouco se tem falado sobre incerteza associada a apreciação.

A proposta deste trabalho é, tendo como base a apreciação individual de diversos campos com as mesmas características (ex: início da produção, tipo de óleo, de campo, etc.) e supondo que para cada ano após o início da produção, a apreciação desses campos nesse ano possui uma distribuição conhecida, determinar a apreciação média e os limites de confiança inferior e superior para cada ano.

Além disso, para se estabelecer os limites de confiança é necessário formular uma hipótese sobre a distribuição das taxas de apreciação.

5.2. Conceitos Estatísticos

Neste item serão abordados alguns conceitos estatísticos, importantes para o estudo.

5.2.1. Teorema Central do Limite

Se uma variável aleatória Y é formada pela soma de variáveis aleatórias independentes, então para um tamanho de m suficientemente grande, tem-se Y normalmente distribuída conforme o teorema a seguir. (Hines & Montgomery, 1990)

Teorema 1

Se X_1, X_2, \dots, X_m é uma seqüência de m variáveis aleatórias independentes com $E(X_j) = \mu_j$ e $V(X_j) = \sigma_j^2$ (ambos finitos) e $Y = X_1 + X_2 + \dots + X_m$, então:

$$Z_m = \frac{Y - \sum_{j=1}^m \mu_j}{\sqrt{\sum_{j=1}^m \sigma_j^2}} \quad (14)$$

Tem uma distribuição aproximada $N(0,1)$, quando n tende para o infinito. Se F_m é a função distribuição de Z_m , então,

$$\lim_{m \rightarrow \infty} \frac{F_m(Z)}{\phi(Z)} = 1 \quad (15)$$

As condições gerais citadas podem ser interpretadas como os termos X_j tomados individualmente, contribuem com um valor desprezível para a variância da soma, e não é provável que um termo possa fazer grandes contribuições na soma.

O uso da distribuição normal é em virtude de que Y tem uma distribuição aproximadamente normal quando a variável X_j pode ter qualquer distribuição.

Tem-se um caso especial da teorema do limite central quando cada componente X_j possui a mesma distribuição.

Teorema 2

Se X_1, X_2, \dots, X_m é uma seqüência de n variáveis aleatórias independentes identicamente distribuídos com $E(X_j) = \mu$ e $V(X_j) = \sigma^2$, e $Y=X_1+X_2+\dots+X_m$, então:

$$Z_m = \frac{Y - m\mu}{\sigma\sqrt{m}} \quad (16)$$

Tem uma distribuição aproximada $N(0,1)$ e se F_m é uma função distribuição de Z_m , então,

$$\lim_{m \rightarrow \infty} \frac{F_m(Z)}{\phi(Z)} = 1 \quad (17)$$

5.2.2. Intervalo de Confiança

O objetivo é estabelecer um IC (intervalo de confiança) para o valor médio (esperado) do fator de apreciação para ano após o início da produção i , ou seja, o limite inferior L e um limite superior U tais que:

$$P\{L \leq \theta \leq U\} \geq 1 - \alpha \quad (18)$$

com intervalo resultante,

$$L \leq \theta \leq U \quad (19)$$

Os valores de L e U podem incluir $-\infty$ e $+\infty$. $1 - \alpha$ é denominado de coeficiente de confiança do intervalo L e U .

Com base em que $Z = \left(\frac{\bar{x} - \mu}{S/\sqrt{m}} \right)$ segue uma distribuição t , com $m - 1$ graus de liberdade, e com

$$\bar{X} = \frac{1}{m} \sum_{j=1}^m x_j \quad (20)$$

$$S^2 = \sum_{j=1}^m \frac{(x_j - \bar{x})^2}{(m-1)} \quad (21)$$

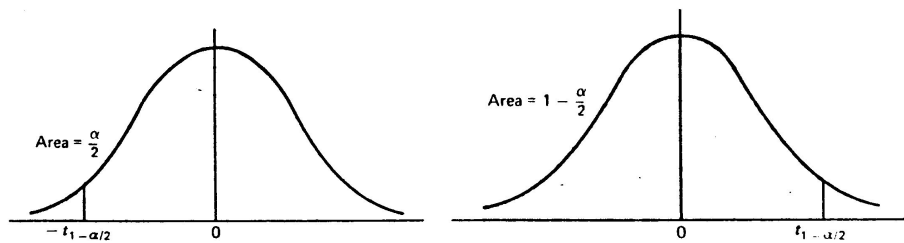


Figura 12 – Distribuição t (Larson, 1982)

Conforme a Figura 12, tem-se,

$$P(t_{-1-\alpha/2} \leq \frac{(\bar{X} - \mu)}{S/\sqrt{m}} \leq t_{1-\alpha/2}) = 1 - \alpha \quad (22)$$

onde $\frac{(\bar{X} - \mu)}{S/\sqrt{m}} \leq t_{1-\alpha/2}$, desenvolvendo, tem-se,

$$\mu \leq \bar{X} + \frac{St_{1-\alpha/2}}{\sqrt{m}} \quad (23)$$

O mesmo desenvolvimento pode ser feito para a outra cauda da curva, obtendo,

$$\left\{ \bar{X} - \frac{St_{1-\alpha/2}}{\sqrt{m}} \leq \mu \leq \bar{X} + \frac{St_{1-\alpha/2}}{\sqrt{m}} \right\} \quad (24)$$

Portanto definem-se os limites inferiores e superiores como:

$$L = \bar{X} - \frac{St_{1-\alpha/2}}{\sqrt{m}} \quad (25)$$

$$U = \bar{X} + \frac{St_{1-\alpha/2}}{\sqrt{m}} \quad (26)$$

onde $t_{1-\alpha/2}$ é o fractil de $1-\alpha/2$ da variável t de *student* com $m-1$ graus de liberdade. O valor de $t_{1-\alpha/2}$ pode ser obtido na tabela da distribuição t no Anexo (Tabela 1).

5.2.3. Determinação dos *Outliers*

Existem diversas formas para determinar *outliers*. Nesse trabalho será utilizado o teste de *Grubbs*, que se baseia nos resíduos padronizados. É um teste unilateral, no qual os dados são dispostos em ordem crescente, de modo visualizar a discrepância de alguns dados quando comparados com os demais. As estatísticas de testes são:

$$G_{\min} = \bar{G} - Grubbs \times S \quad (27)$$

$$G_{\max} = Grubbs \times S - \bar{G} \quad (28)$$

onde G_{\min} e G_{\max} são suspeitos de serem *outliers*

\bar{G} é o valor médio das observações amostrais

S é o desvio-padrão da amostra

As estatísticas *Grubbs* determinados como valores críticos tabelados segundo o nível de significância e o tamanho m da amostra. Os valores das estatísticas *Grubbs* são apresentados no quadro a seguir para um nível de significância de 95%.

Quadro 3 – Valor da estatística *Grubbs* para uma amostra de tamanho *m*

<i>m</i>	<i>Grubbs</i>	<i>m</i>	<i>Grubbs</i>	<i>m</i>	<i>Grubbs</i>	<i>m</i>	<i>Grubbs</i>	<i>m</i>	<i>Grubbs</i>
3	1.15	13	2.46	23	2.78	33	2.95	70	3.26
4	1.48	14	2.51	24	2.80	34	2.97	80	3.31
5	1.71	15	2.55	25	2.82	35	2.98	90	3.35
6	1.89	16	2.59	26	2.84	36	2.99	100	3.38
7	2.02	17	2.62	27	2.86	37	3.00	110	3.42
8	2.13	18	2.65	28	2.88	38	3.01	120	3.44
9	2.21	19	2.68	29	2.89	39	3.03	130	3.47
10	2.29	20	2.71	30	2.91	40	3.04	140	3.49
11	2.34	21	2.73	31	2.92	50	3.13		
12	2.41	22	2.76	32	2.94	60	3.20		

Fonte: Graphpad, 2003.

5.3.

Determinação do intervalo de Confiança da Avaliação de Reservas Média

Para se determinar a avaliação é necessário obter dados de crescimento de reservas provadas de uma amostra de *m* campos de características homogêneas desde início da produção até um dado período de tempo desde o início da produção dos campos.

Aconselha-se o agrupamento por características semelhantes conforme citado no Capítulo 4 e de tamanho por Watkins (2000). Supõe-se que a agregação de mesmas características faça com que os vários campos utilizados como base tenham comportamento e valores de avaliação muito próximos uns dos outros, podendo seu padrão comum de crescimento servir como base para inferência sobre o crescimento de novos campos similares.

Deve-se determinar os fatores de avaliação cumulativos para todos os campos para cada ano *i* utilizando a fórmula $G_{ij} = \frac{\text{Reserva Provada}_{ij}}{\text{Reserva Provada}_{0j}}$, onde G_{ij} é

o *i*-ésimo fator de avaliação para o *j*-ésimo campo, $\text{Reserva Provada}_{ij}$ é a reserva provada de óleo ou gás para o *i*-ésimo ano após o início da produção do *j*-ésimo campo e $\text{Reserva Provada}_{0j}$ é a reserva provada de óleo ou gás quando iniciou-se a produção do *j*-ésimo campo. ($i=0, 1, 2, \dots, n$ e $j=1, 2, \dots, m$)

Quadro 4 – Tabela de Apreciação de Reservas dos m Campos

Ano(i)	Campo 1	Campo 2	Campo m
0	G_{01}	G_{02}	G_{0m}
1	G_{11}	G_{12}	G_{1m}
2	G_{21}	G_{22}	G_{2m}
\vdots	\vdots	\vdots	\ddots	\vdots
n	G_{n1}	G_{n2}	G_{nm}

O Quadro 4 mostra um exemplo de resumo da apreciação de reservas para n anos após o início da produção de m campos.

Após determinar os fatores de apreciação de reservas cumulativos para n anos dos m campos, determinam-se a média, a variância e o desvio-padrão das variáveis de apreciação de cada campo no i -ésimo ano. As estimativas da média e da variância podem ser obtidas pelas seguintes equações:

$$\bar{G}_i = \frac{1}{m} \sum_{j=1}^m G_{ij} \quad (29)$$

$$S_i^2 = \sum_{j=1}^m \frac{(G_{ij} - \bar{G}_i)^2}{(m-1)} \quad (30)$$

e, conseqüentemente, a estimativa do desvio-padrão é dada por,

$$S_i = \sqrt{S_i^2} \quad (31)$$

onde \bar{G}_i é a apreciação média para o i -ésimo ano, S_i^2 é a variância para o i -ésimo ano e S_i é o desvio padrão para i -ésimo ano. Todos os valores de média e variância para o i -ésimo ano são determinados utilizando as m variáveis de apreciação para o ano em questão.

Após determinar as estimativas dos parâmetros estatísticos, calcula-se os limites inferior e superior em relação à média para cada i -ésimo ano após o início da produção. Os limites inferiores e superiores são obtidos com base nos parâmetros determinados anteriormente (média e desvio-padrão), pelas seguintes equações.

$$LI_i = \bar{G}_i - \frac{S_i t_{1-\alpha/2}}{\sqrt{n}} \quad (32)$$

$$LS_i = \bar{G}_i + \frac{S_i t_{1-\alpha/2}}{\sqrt{n}} \quad (33)$$

Onde LI_i é o limite inferior em relação à média para o i -ésimo ano e LS_i é o limite superior em relação à média para o i -ésimo ano. $t_{1-\alpha/2}$ é o fractil de $1-\alpha/2$ da variável t de *student* com $m-1$ graus de liberdade (valor de t o qual $P(T > t) = \alpha/2$). O valor de $t_{1-\alpha/2}$ pode ser obtido na Tabela 1 da distribuição t no Anexo.

Os valores da média, limites inferiores e superiores da média para ano i após o início de produção, podem ser resumidos como na tabela exemplo do Quadro 5.

Quadro 5 – Tabela Resumo de média e limites inferiores e superiores de cada ano i

Ano (i)	LI_i	Média _{i}	LS_i
1	LI_1	Média ₁	LS_1
2	LI_2	Média ₂	LS_2
3	LI_3	Média ₃	LS_3
\vdots	\vdots	\vdots	\vdots
n	LI_n	Média _n	LS_n

Determinar os possíveis campos *outliers* da amostra por meio do teste de *Grubbs* representados pelas estatísticas:

$$G_{\min} = \bar{G} - Grubbs \times S \quad (34)$$

$$G_{\max} = Grubbs \times S - \bar{G} \quad (35)$$

Após determinar os valores máximos e mínimos da apreciação de reservas para cada ano após o início da produção, compará-los com os dados de apreciação de reservas dos campos analisados, determinando, dessa forma, os *outliers*.

5.4.

Exemplo de Cálculo do Intervalo de Confiança da Avaliação Média de Reservas com 5 Campos de Óleo Onshore

Para exemplificar a metodologia de avaliação de reservas cujo crescimento é sujeito a uma tendência logarítmica e com intervalos de confiança para cada ano após o início da produção foi utilizado o conjunto de dados de cinco campos *onshore*. Apesar de a empresa fornecedora dos dados ter considerado a amostra homogênea, as informações sobre as características dos mesmos não foi fornecidas por questão de confidencialidade, de modo que a verificação de semelhança das características não pode ser melhor estudada. É importante salientar que o estudo deveria ter sido feito com um número maior de campos ($m \geq 15$) para maior significância dos estimadores. Além disso, com esse maior número de campos a hipótese de distribuição normal dos fatores de apreciação teria apoio no teorema central do limite.

A Figura a seguir mostra a planilha eletrônica onde foram calculadas as avaliações.

	A	B	C	D	E	F	G	H	I	J
1	Ano(i)	Campo1	Campo2	Campo3	Campo4	Campo5		Media _i	Var _i	Dev.Pad _i
2	0	1	1	1	1	1		1	0	0
3	1	5,340797	2,125111	0,97886	1,096988	1,652059		2,23876277	3,217689998	1,793792072
4	2	5,44333	3,107133	1,013313	0,892586	3,560982		2,803468902	3,622759703	1,903354855
5	3	5,608919	3,777688	1,031158	2,563185	3,347293		3,265648611	2,81154279	1,676765574
6	4	5,158715	3,840074	1,568876	3,326352	3,905833		3,559969883	1,69327477	1,301258917
7	5	5,66108	4,5269	1,859293	2,691921	4,629946		3,873828213	2,411935577	1,553040752
8	6	6,281137	6,791264	1,837813	2,889683	4,860285		4,532036297	4,556843211	2,134676372
9	7	5,970489	6,990294	1,791142	2,882759	6,515047		4,82994622	5,4583429	2,336309676
10	8	5,921666	6,936723	1,764744	3,162276	5,876741		4,73242994	4,713795163	2,171127625
11	9	6,088754	6,728163	1,813392	3,415722	5,889324		4,787070943	4,35013598	2,08569796
12										
13	n =	10								
14										
15	$t_{1-\alpha/2}$ =	2,776								
16										
17		Ano(i)	LI _i	Media _i	LS _i		Ano(i)	LI _i	Media _i	LS _i
18		0	1	1	1		0	I18-(J2*B15)/RAIZ(B13)	H2	I18+(J2*B15)/RAIZ(B13)
19		1	0,664085	2,238763	3,81344		1	I19-(J3*B15)/RAIZ(B13)	H3	I19+(J3*B15)/RAIZ(B13)
20		2	1,132612	2,803469	4,474326		2	I20-(J4*B15)/RAIZ(B13)	H4	I20+(J4*B15)/RAIZ(B13)
21		3	1,793703	3,265649	4,737594		3	I21-(J5*B15)/RAIZ(B13)	H5	I21+(J5*B15)/RAIZ(B13)
22		4	2,417662	3,55997	4,702278		4	I22-(J6*B15)/RAIZ(B13)	H6	I22+(J6*B15)/RAIZ(B13)
23		5	2,510494	3,873828	5,237162		5	I23-(J7*B15)/RAIZ(B13)	H7	I23+(J7*B15)/RAIZ(B13)
24		6	2,658114	4,532036	6,405958		6	I24-(J8*B15)/RAIZ(B13)	H8	I24+(J8*B15)/RAIZ(B13)
25		7	2,779021	4,829946	6,880872		7	I25-(J9*B15)/RAIZ(B13)	H9	I25+(J9*B15)/RAIZ(B13)
26		8	2,826509	4,73243	6,638351		8	I26-(J10*B15)/RAIZ(B13)	H10	I26+(J10*B15)/RAIZ(B13)
27		9	2,956145	4,787071	6,617997		9	I27-(J11*B15)/RAIZ(B13)	H11	I27+(J11*B15)/RAIZ(B13)

Figura 13 – Determinação dos Limites inferiores e superiores e médias

Com base nas avaliações de reservas dos cinco campos, determinaram-se a média e o desvio-padrão para cada *i-ésimo* ano após o início da produção. Com base nesses parâmetros calculados, determinou-se o intervalo de confiança, representado pelo limite inferior e superior, para cada ano após início da produção. A seguir plotou-se o intervalo de confiança e o valor médio da

apreciação para cada ano após o início da produção juntamente com a apreciação dos cinco campos analisados.

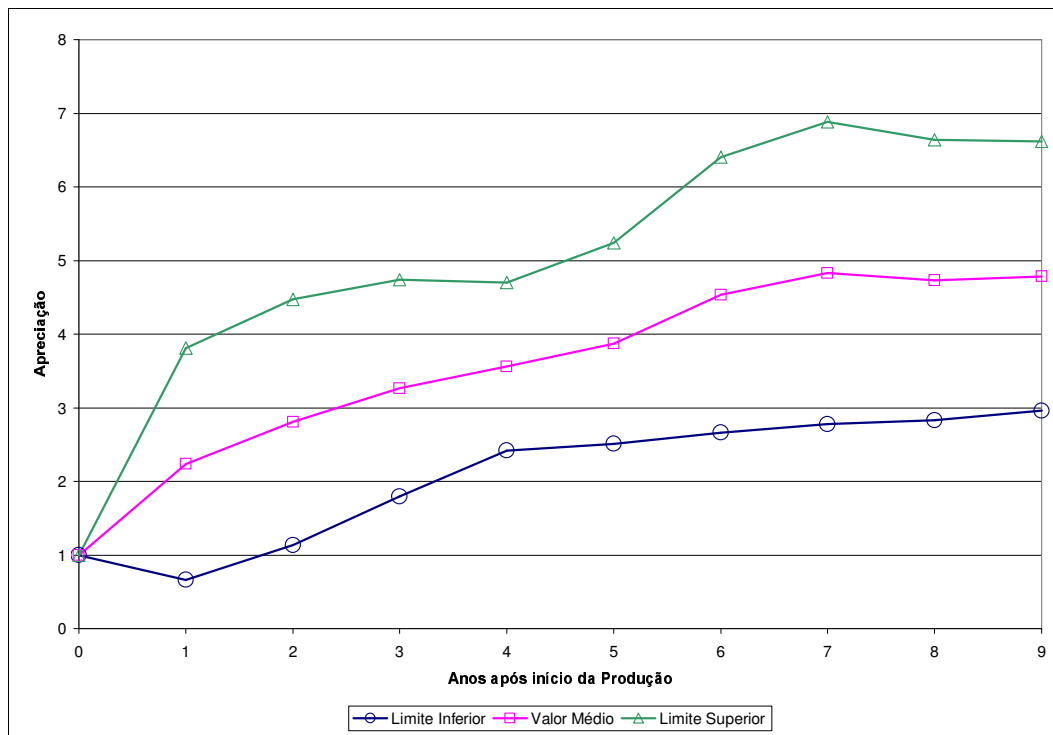


Figura 14 – Intervalo de confiança e valor médio

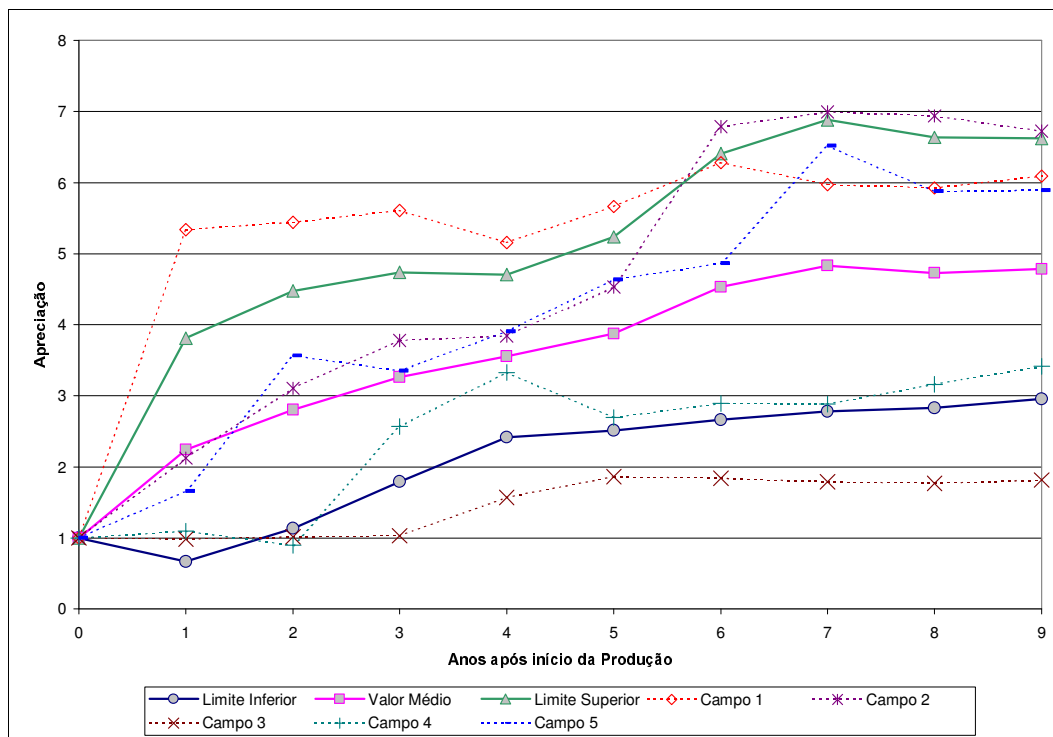


Figura 15 – Intervalo de confiança, valor médio e apreciação dos campos

Os possíveis *outliers* foram analisados por meio do teste de *Grubbs*, como pode ser observado na Figura 16 a seguir.

Ano(i)	G _i max	G _i min			
0	1	1	Grubbs para N=5 1.71		
1	5.306147	-0.828622			
2	6.058206	-0.451268			
3	6.132918	0.398379			
4	5.785123	1.334817			
5	6.529528	1.218129			
6	8.182333	0.88174			
7	8.825036	0.834857			
8	8.445058	1.019802			
9	8.353614	1.220527			
Ano(i)	Campo1	Campo2	Campo3	Campo4	Campo5
0	1	1	1	1	1
1	5.340797	2.125111	0.97886	1.096988	1.652059
2	5.44333	3.107133	1.013313	0.892586	3.560982
3	5.608919	3.777688	1.031158	2.563185	3.347293
4	5.158715	3.840074	1.568876	3.326352	3.905833
5	5.66108	4.5269	1.859293	2.691921	4.629946
6	6.281137	6.791264	1.837813	2.889683	4.860285
7	5.970489	6.990294	1.791142	2.882759	6.515047
8	5.921666	6.936723	1.764744	3.162276	5.876741
9	6.088754	6.728163	1.813392	3.415722	5.889324

Figura 16 – Teste de *Grubbs*

Pela Figura observa-se que apenas o campo 1, representado pelo primeiro ano após o início da produção se trata de um *outlier*, e portanto esse, deveria ser retirado da amostra.

Poder-se-ia ajustar uma forma funcional para o valor médio, entretanto observa-se que as apreciações de reservas desses campos analisadas apresentam um caso clássico de heteroscedaticidade, como pode ser observado na Figura a seguir.

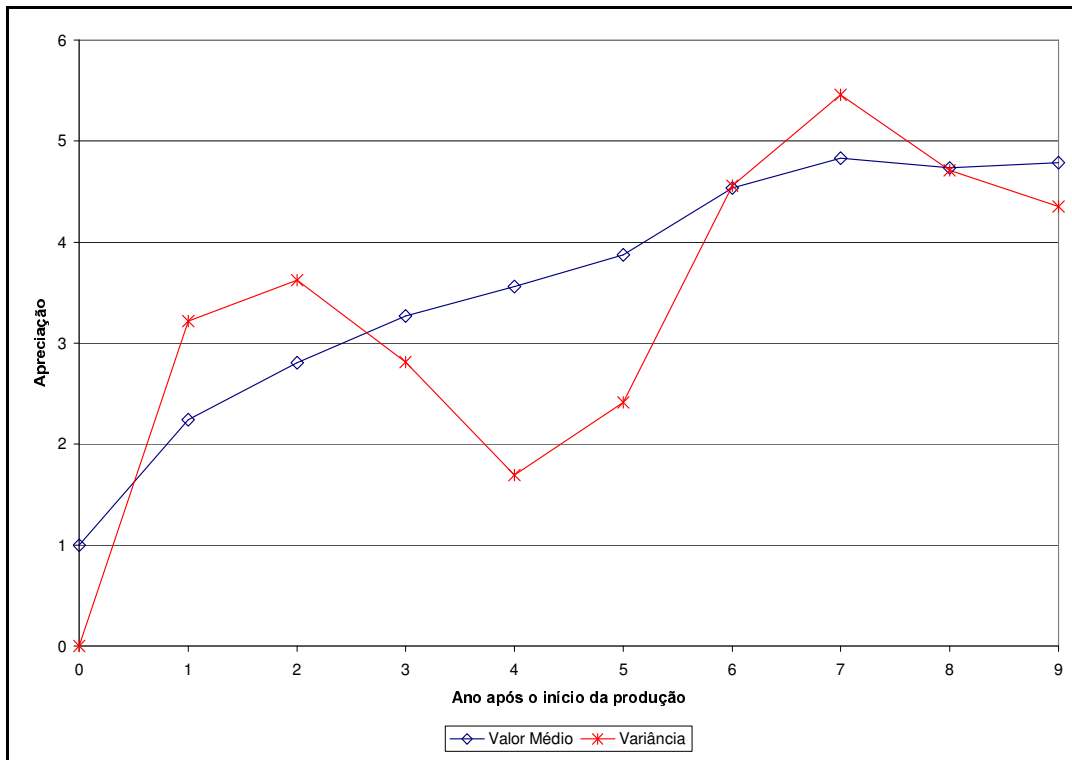


Figura 17 – Valor médio e Variância

Poderia-se fazer uma transformação para corrigir essa variância irregular, entretanto, nesse estudo não será possível essa transformação, pois a variância não segue uma tendência regular. No Anexo têm-se os procedimentos da regressão para ajustar uma função logarítmica, quando a variância da apreciação for uniforme.