

Cálculo de Factores de Expansão para os Dados do Trabalho de Inquérito Agrícola 2002

David J. Megill
Consultor de Amostragem

Dezembro de 2002

1. Introdução

O desenho da amostra para o Trabalho de Inquérito Agrícola (TIA) 2002 está documentado no relatório “Análise de Erros Padrão Simulados para os Resultados do Trabalho de Inquérito Agrícola 2002 Baseados numa Amostra de 80 Distritos para Moçambique” (Megill, Abril de 2002). O Ministério de Agricultura e Desenvolvimento Rural (MADER) fez o trabalho de campo para este inquérito entre julho e outubro de 2002, em 80 distritos seleccionados para o TIA 2002. O propósito de este relatório é de especificar os procedimentos usados para calcular os factores de expansão para os dados do inquérito. O desenho da amostra para o TIA 2002 também está sumariado aqui, dado que os factores de expansão são baseados na metodologia da amostragem.

2. Desenho da Amostra do TIA 2002

O Censo Agro-Pecuário (CAP) foi usado como base de amostragem para o TIA 2002. Uma sub-amostra das unidades primárias de amostragem (UPAs) ou segmentos do CAP foram seleccionadas para o TIA. A amostra está estratificada por província e zona ecológica (26 estratos). Em 18 estratos a amostra do TIA foi seleccionada em três etapas de amostragem: os distritos foram seleccionados na primeira etapa; sete segmentos foram seleccionados em cada distrito amostral na segunda etapa; e agregados familiares foram seleccionados na terceira etapa. Nos outros 8 estratos, a selecção foi bi-etápica, dado que os segmentos foram seleccionados na primeira etapa. Nestes estratos bi-etápicas o número de segmentos seleccionados por distrito varia com o tamanho do distrito, com um promédio de 7 segmentos amostrais por distrito. Depois da listagem de agregados familiares dentro de cada segmento amostral, 8 pequenas explorações foram seleccionadas por segmento na última etapa. A atribuição dos 80 distritos amostrais, e o número de segmentos e pequenas explorações na amostra do TIA 2002 por estrato está apresentada na Tabela 1.

Tabela 1. Atribuição da Amostra por Estrato para o TIA 2002 com 80 Distritos Amostrais:
Distritos, Segmentos e Pequenas Explorações

Província	Estrato	No. Distritos CAP	No. Distritos TIA	No. Segmentos TIA	No. Pequenas Explorações TIA
Niassa	0101	5	2	14	112
	0106	11	3	21	168
Cabo Delgado	0205	2	2	14	112
	0206	2	2	14	112
	0210	13	5	35	280
Nampula	0306	3	2	14	112
	0310	15	7	49	392
	0311	2	2	14	112
Zambézia	0404	3	3	21	168
	0407	5	5	35	280
	0411	9	5	35	280
Tete	0502	3	3	21	168
	0508	3	2	14	112
	0512	7	3	21	168
Manica	0603	5	3	21	168
	0609	3	3	21	168
	0614	2	2	14	112
Sofala	0709	4	2	14	112
	0711	5	3	21	168
	0712	3	2	14	112
Inhambane	0813	11	5	35	280
	0814	3	2	14	112
Gaza	0913	3	2	14	112
	0914	5	3	21	168
	0915	4	2	14	112
Maputo- Província	1014	7	5	35	280
Moçambique		138	80	560	4,480

A amostra do TIA 2002 também foi estratificada por tamanho da exploração, usando os critérios especificados abaixo na Tabela 2.

Tabela 2: *Classificação de explorações*

Factores	Limite 1	Limite 2
Área cultivada não irrigada (Ha)	10	50
Número de cabeças de Gado Bovino	10	100
Número de Caprinos/Ovinos/Suínos	50	500
Número de aves ¹⁾	5.000	20.000

¹⁾ Em relação às aves, a exploração para ser considerada média ou grande, deve exercer a actividade de forma contínua, pelo menos nos últimos 6 meses.

Se todos os factores forem menores que limite 1 a exploração é pequena;

Se um factor for maior ou igual a valores do limite 1 e menor que o limite 2 a exploração passa para a categoria de exploração média;

Se um factor for maior ou igual a valores do limite 2, passa para categoria de grande exploração.

Nota1: As explorações podem ser **agrícolas** (quando se dedicam somente a produção de culturas), **Pecuárias** (criação de animais) ou apresentam-se de forma combinada.

Nota 2: Área cultivada consiste de área com culturas anuais mais área em pousio, não incluindo área em pastagem ou com culturas permanentes.

Nota 3: Para os casos das **hortícolas, pomares e plantações produtivas e áreas irrigadas** a sua classificação obedece a seguinte distribuição:

- a) Se a área for **menor ou igual a 5 ha** deve ser considerada pequena exploração;
- b) Se a área estiver **entre 5 e 10 ha** deve ser considerada média exploração;
- c) Se a área for **maior que 10 ha** deve ser considerada grande exploração.

Source: TIA 2002. Manual do Inquiridor.

Todas as médias e grandes explorações listadas dentro dos segmentos seleccionados foram incluídas na amostra do TIA. Por isso as médias explorações têm a mesma probabilidade de selecção do segmento amostral. Todas as grandes explorações identificadas na base para o distrito foram incluídas na amostra do TIA, e têm a mesma probabilidade de selecção do distrito. No caso dos estratos com amostragem bi-etápica, as grandes explorações são auto-representadas, quer dizer, têm probabilidade igual a 1.

3. Procedimentos Usados para Calcular os Factores de Expansão para o TIA 2002

Para que as estimativas de um inquérito sejam representativas da população, é necessário multiplicar os dados por um factor de expansão. O factor de expansão básico para cada agregado familiar seleccionado seria igual ao inverso de sua probabilidade de selecção (calculado multiplicando as probabilidades para cada etapa de amostragem). Dado que as probabilidades variam por estrato de acordo ao tamanho da exploração, os factores de expansão estão especificados aqui individualmente para cada tamanho de exploração.

3.1. Factor de Expansão para as Grandes Explorações

No caso das grandes explorações na base do distrito, a probabilidade de selecção seria igual à probabilidade do distrito. Os distritos dentro dos estratos com amostragem tri-etápica foram seleccionados com probabilidades proporcional ao tamanho, baseados no número de agregados familiares na base de amostragem do CAP, gerado dos dados do Recenseamento Geral da População e Habitação (RGPH) de 1997. O factor de expansão para as grandes explorações seria calculado como o inverso desta probabilidade:

$$W_{hi(g)} = \frac{M_h}{n_h \times M_{hi}}$$

onde:

$W_{hi(g)}$ = factor de expansão básico para as grandes explorações no distrito amostral i do estrato h; nos estratos com amostragem bi-etápica, $W_{hi(g)} = 1$

M_h = número total de agregados familiares dentro do estrato h na base de amostragem do CAP

n_h = número de distritos seleccionados para o TIA 2002 no estrato h

M_{hi} = número total de agregados familiares dentro do distrito i do estrato h na base de amostragem do CAP

É importante ajustar os factores de expansão para tomar em conta os agregados familiares não inqueridos. O factor de expansão final ($W'_{hi(g)}$) para as grandes explorações pode ser expressada da seguinte maneira:

$$W'_{hi(g)} = W_{hi(g)} \times \frac{M_{hi(g)}}{M'_{hi(g)}}$$

onde:

$M_{hi(g)}$ = número total de grandes explorações na base para o distrito i do estrato h

$M'_{hi(g)}$ = número de grandes explorações com questionários do TIA completados no distrito i do estrato h

Na listagem dentro dos segmentos seleccionados para o TIA 2002, outras explorações foram

identificadas com os critérios de grandes, mas é importante distinguir estas explorações daquelas na base original de grandes explorações do distrito, porque as probabilidades são diferentes. No caso das grandes explorações identificadas na listagem do segmento amostral (e que foram omitidas na base de grandes explorações do distrito), o factor de expansão seria o mesmo que aquele especificado abaixo para as médias explorações.

3.2. Factor de Expansão para as Médias Explorações

No caso das médias explorações, a probabilidade de selecção seria igual à probabilidade do segmento amostral, baseada na probabilidade de selecção da área de enumeração (AE) correspondente para o CAP; esta probabilidade foi proporcional ao tamanho (número de agregados familiares). Dado a metodologia de amostragem para o CAP, não podemos calcular a probabilidade exacta da selecção de cada segmento, mas podemos supor que foi seleccionado com probabilidade proporcional ao número de agregados familiares. Neste caso a melhor fonte de dados para a medida de tamanho seria a listagem no segmento amostral para o TIA. Dado que esta medida seria o número actual de agregados familiares no segmento (e não o tamanho original na base de amostragem), também devemos usar o número actual de agregados familiares no distrito, estimado das projecções demográficas, para calcular o factor de expansão. O factor de expansão para as médias explorações seria calculado da seguinte maneira:

$$W_{hij(m)} = \frac{\hat{M}_h}{7 \times n_h \times M_{hij}},$$

onde:

$W_{hij(m)}$ = factor de expansão básico para as médias explorações no segmento j do distrito amostral i no estrato h

\hat{M}_h = número total de agregados familiares dentro do estrato h para a metade do ano 2002, estimado das projecções demográficas baseadas no RGPH de 1997

n_h = número de distritos seleccionados para o TIA 2002 no estrato h

M_{hij} = número de agregados familiares listados no segmento amostral j do distrito amostral i no estrato h

No caso dos estratos com amostragem bi-etápica, o número de segmentos seleccionados também seria igual a $7 \times n_h$; então esta expressão também pode ser usada para calcular os factores de expansão das médias explorações nestes estratos. As projecções demográficas disponíveis tem uma data de referencia que corresponde à metade do ano 2002 (30 de junho), que é perto da metade do período do trabalho de campo para o TIA 2002.

No caso do estrato 0609 em Manica, não foi possível fazer inquérito em um dos segmentos

seleccionados no distrito de Mossurize. Dado que a amostragem foi bi-etápica neste estrato com 21 segmentos seleccionados, o peso básico foi ajustado por um factor de 21/20.

Seria necessário ajustar estes factores de expansão para tomar em conta os agregados familiares não inqueridos. O factor de expansão final ($W'_{hij(m)}$) para as médias explorações pode ser expressada da seguinte maneira:

$$W'_{hij(m)} = W_{hij(m)} \times \frac{M_{hij(m)}}{M'_{hij(m)}},$$

onde:

$M_{hij(m)}$ = número de médias (e grandes) explorações listadas no segmento amostral j do distrito amostral i no estrato h

$M'_{hij(m)}$ = número de médias (e grandes) explorações com questionários do TIA completados no segmento amostral j do distrito amostral i no estrato h

3.3. Factor de Expansão para as Pequenas Explorações

A amostra para as pequenas explorações foi desenhada para ser aproximadamente auto-ponderada (com probabilidades iguais) dentro de cada estrato. Então usamos a seguinte fórmula para estimar o factor de expansão para as pequenas explorações:

$$W_{h(p)} = \frac{\hat{M}_h \& \hat{M}_{h(mg)}}{m_{h(p)}},$$

onde:

$W_{h(p)}$ = factor de expansão para as pequenas explorações seleccionadas no estrato h para o TIA 2002

$\hat{M}_{h(mg)}$ = estimativa do número total de médias e grandes explorações no estrato h, baseado nos dados ponderados da listagem

$m_{h(p)} = 56 \times n_h$ = número de pequenas explorações seleccionadas no estrato h para o TIA 2002

No caso do estrato 0609 em Manica, foi necessário aplicar um factor de ajuste de 21/20 para

tomar em conta um segmento seleccionado que não foi inquerido.

Como no caso das médias explorações, também é recomendável ajustar os factores de expansão para as pequenas explorações a nível de segmento para tomar em conta os agregados familiares não inqueridos. O factor de expansão final ($W'_{hij(p)}$) para as pequenas explorações dentro de cada segmento amostral pode ser expressada da seguinte maneira:

$$W'_{hij(p)} = W_{h(p)} \times \frac{m_{hij(p)}}{m_{hij(p)}},$$

onde:

$m_{hij(p)} =$ = 8 = número de pequenas explorações seleccionadas no segmento amostral j do distrito amostral i no estrato h

$m'_{hij(p)} =$ número de pequenas explorações com questionários do TIA completados no segmento amostral j do distrito amostral i no estrato h