

INSTITUTO NACIONAL DE ESTATÍSTICA

ASPECTOS METODOLÓGICOS

MOVIMENTO DE PESSOAS NAS FRONTEIRAS

Versão 1

INDICE

INTRODUÇÃO.....	3
1-ÂMBITO GERAL DO ESTUDO	3
1.1 Âmbito geográfico/domínio de estudo	4
1.2 Fonte de informação utilizada.....	5
1.3 Unidades amostrais e unidades de observação	5
1.4- Descrição do método de recolha	5
1.5 Confidencialidade dos dados.....	6
1.6 Avaliação da Qualidade Estatística e harmonização.....	6
1.7- Tratamento dos Dados	6
1.8 Periodicidade de realização da operação em 2011.....	7
2- PLANO DE AMOSTRAGEM.....	8
2.1-Tamanho e estrutura da amostra.....	9
2.2- Base de Amostragem e População Alvo.....	15
3-ALGUNS ASPECTOS OPERACIONAIS DO INQUÉRITO.....	17
3.1 Selecção dos indivíduos: amostra inicial.....	17
3.2- Como seleccionar um turista?	19
3.3 Observações importantes: resumo.....	22
4- DETERMINAÇÃO DA AMOSTRA FINAL.....	23
4.1-Amostra final de cada domínio	23
4.2-Amostra final Global	23
4.3 Ponderação à posteriori das proporções de turistas não residentes	23
4.3.1-Cálculo do factor de ponderação ou de expansão	24
ANEXO A: OUTROS CONCEITOS IMPORTANTES.....	25
ANEXO B: RESULTADOS (A DESENVOLVER)!!	Erro! Marcador não definido.

INTRODUÇÃO

O presente documento descreve, em grandes linhas, os diferentes aspectos metodológicos que estão subjacentes à realização do Inquérito à Satisfação e Gastos dos Turistas (ISGT) de 2011. Pretende-se também descrever e documentar as diferentes opções metodologias saídas das reuniões de discussão e de socialização realizadas no Instituto Nacional de Estatística, com intuito de capitalizar as experiências adquiridas na realização de vários inquéritos demográficos e socioeconómicos em prol da realização, com êxito, do ISGT.

1-ÂMBITO GERAL DO ESTUDO

O Inquérito à Satisfação e Gastos dos Turistas é uma operação estatística que visa inquirir uma amostra de visitantes¹, nomeadamente, sobre os gastos realizado e à satisfação pelos serviços que lhes são prestados em Cabo Verde. Tendo em conta o fluxo de passageiros e a frequência crescente de voos internacionais que acontecem nesses aeroportos, principalmente, numa conjuntura em que turismo constitui um dos sectores com maior dinâmica de desenvolvimento económico e social, é fundamental obter dados actualizados sobre os visitantes na medida em que eles contribuem consideravelmente para a entrada de divisas, para a promoção de emprego e ao mesmo tempo para a redução da pobreza.

O presente plano de amostragem procura ser eficaz ao nível de precisão desejável, no sentido de satisfazer as seguintes OBJECTIVOS ao nível nacional, nomeadamente:

Estimar alguns parâmetros de interesse (valores desconhecidos, mas fixo numa população), tais como:

- 1) A Média e o total de gastos realizados pelos turistas em Cabo verde;
- 2) A Percentagem de turistas que estão satisfeitos com a qualidade e preço dos serviços;
- 3) A percentagem de turistas com gasto superior a um determinado valor;

¹ “A Organização Mundial de Turismo (OMT) define como visitante TODO E QUALQUER individuo que viaja a um local que esteja fora do seu ambiente habitual, por um período inferior a 12 meses e cujo motivo principal da visita não seja o de exercer uma actividade remunerada no local visitado”.

Aferir sobre a estrutura de gasto turístico, a fim de estabelecer um perfil actualizado do turismo em Cabo Verde e obter elementos para a montagem de um sistema de cálculo da conta satélite do turismo.

1.1 Âmbito geográfico/domínio de estudo

Este inquérito por amostragem realiza-se nos Aeroportos Internacionais de Cabo Verde, no momento de saída do país, tanto para os não residentes como para residentes nos seguintes domínios:

- 1) Aeroporto internacional da Praia
- 2) Aeroporto internacional da Sal
- 3) Aeroporto internacional da Boa Vista
- 4) Aeroporto internacional da S.Vicente

O fluxo de visitantes nestes três primeiros aeroportos representa a grande maioria (cerca de 98 %) dos movimentos dos hóspedes nos diferentes hotéis do país FONTE: INECV- (DATA)???. Os visitantes que afluem a estes aeroportos são basicamente os que preferem um turismo de “sol e praia”. Por sua vez, os visitantes que afluem ao Aeroporto Internacional de S.Vicente, mesmo não tendo expressividade nacional quanto à contribuição no fluxo turístico (inferior a 5%: com cerca de 1,7 %), apresentam um perfil turístico diferenciado caracterizado como turismo “rural e de montanhas” contrapondo assim ao turismo de “sol e praia” como supracitados. Sendo assim, apesar da abertura do aeroporto Internacional de S.Vicente ser recente, graça ao perfil diferenciado do tipo de turismo que se verifica neste aeroporto e à tendência crescente do fluxo de visitantes observados no mesmo aeroporto, a inclusão do Aeroporto Internacional de S.Vicente como um dos domínios de estudo (todos independentes do ponto de vista da construção da amostra), pode proporcionar maior heterogeneidade à amostra final de turistas. Por conseguinte, pode contribuir para o aumento da representatividade do carácter estatístico da pesquisa do ISGT que ora se realiza. Deste ponto de vista, a avaliação do custo e benefício estatístico pode ser vantajosa se se contemplar este domínio de estudo com intuito de “espelhar” melhor o perfil de turistas que visitam o país, podendo até melhorar a precisão na estimação dos gastos.

1.2 Fonte de informação utilizada

- **Directa** para a componente dos inquiridos por amostragem (**os próprios visitantes**);
- **Indirecta** através das **fronteiras aéreas** para a componente das contagens universais servindo-se destes dados para o ajustamento da amostra e para o cálculo dos factores de ponderação dos elementos da amostra. As contagens indirectas universais são da responsabilidade da Agencia de Segurança Aeroportuária (ASA) /Transporte Aéreo de Cabo Verde (TACV) que disponibilizam a seguinte informação: **Número total de passageiros embarcados: por mês**, por aeroporto, por voo (por exemplo, Transportadora Aéreo Portuguesa (TAP), TACV,..) e por rota (país **de destino**) e **se possível por tipo de voo (voo charter e voo Regular)**, segundo exemplificado no Quadro 1 em anexo.

1.3 Unidades amostrais e unidades de observação

- **A população alvo** é constituída pelos passageiros, embarcados em voos internacionais nos Aeroportos nacionais (residentes e não residentes em território nacional).
- **Unidade de Observação:** Indivíduo;
- **Unidade amostral:** também individuo - quer dizer selecciona-se os indivíduos e observa-se as suas características em todos os voos (cada voo pode ser considerado um conglomerado).

1.4- Descrição do método de recolha

- **Contacto inicial:** não há contacto inicial com o entrevistado.
- **Método de recolha:** auto-preenchimento por parte dos passageiros e sempre que necessário, auxiliado pelos inquiridores.
- **Tipo de dados recolhidos:** _quantitativos e qualitativos referentes às características à caracterização do voo (data de **chegada**, **tipo de viagem**, destino de voo etc) referente à caracterização do turista, questões de opiniões por parte dos turistas etc.,
- **Período de referência dos dados:** mês de inquirição.

- **Período de recolha:** durante 7 dias.
- **Formação aos entrevistadores:** formação teórica e prática
- **Perfil do entrevistador:** Ter pelo menos 12º ano de escolaridade ou equivalente e/ou domínio da língua Francesa e Inglesa. Manual do entrevistador

1.5 Confidencialidade dos dados

- A confidencialidade dos dados está garantida dado que na fase recolha não se procede à identificação do entrevistado.

1.6 Avaliação da Qualidade Estatística e harmonização

- 1) Realizada pela aplicação seguimento das directrizes do plano de sondagem, amostra, amostragem e selecção dos turistas.
- 2) Seguimento escrupuloso das instruções descritas no manual de instruções ou procedimentos dos inquiridores
- 3) Recomendações Nacionais e Internacionais: seguiu-se as recomendações das NU/OMT

1.7- Tratamento dos Dados

- Entrada de dados: digitação – questionário em papel.
- Codificação: manual.
- *Software* utilizado: aplicação informática desenvolvido pelo INE
- Os dados recolhidos serão introduzidos num ficheiro que incluirá todas as variáveis.
- Na fase de digitação estão previstas validações pelo confronto de variáveis relativas à caracterização do voo. As validações referidas procuram, acima de tudo, certificar que a informação observada condiz com o definido na amostra, nomeadamente o tipo de voo, ou o país de destino do voo.
- A informação será analisada através de medidas de estatística descritiva, produção de quadros e gráficos de análise.
- Tratamento de Não Respostas: Não se aplica qualquer tratamento sofisticado para além das imputações dedutivas.

1.8 Periodicidade de realização da operação em 2011

- Época baixa:

Seleccção do mês da época. O mês escolhido será o de OUTUBRO. Este mês foi escolhido a partir de um registo de fluxo, em série, de uma variável que totaliza o número de pessoas entradas nos hotéis por mês ordenados por valores). Verificou-se a evolução desta variável e escolheu-se o mês de Setembro por ser o mês que “melhor” se representava um conjunto de seis meses considerados como os meses da “época baixa”. Na verdade existia um outro mês que correspondia ao menor número de fluxo, mas por ser o mais baixo postulou-se que os preços podiam estar relativamente baixos e os gastos (como uma das principais variáveis podiam estar subvalorizados). Por isso escolheu-se o mês que fosse um dos meses da época baixa e que em teoria não estaria influenciada por esta sazonalidade.

Seleccção da semana: a semana seleccionada será a SEGUNDA semana do mês de OUTUBRO, por ser semana mais estável (a primeira estaria relativamente próxima de Agosto que é considerado como um dos meses da época alta a segunda estaria reservada ao inquérito piloto, e seleccção dos inquiridores).

- Época alta:

Seleccção do mês da época.

O mês escolhido será o de Dezembro. Este mês foi escolhido seguindo o mesmo procedimento metodológico, fazendo referência em relação à época alta. O mês de Dezembro também não é o mês em que se registou maior fluxo de turistas (pois neste mês, ao contrario da época baixa os preços podem estar, em teoria sobreavaliados e a estimação dos gastos tomando este mês como referencia estaria sobrestimados). Mas, será considerado o mês como representativo do conjunto de seis meses considerados como os meses da época alta, e espera-se que em teoria não estaria influenciada por esta sazonalidade.

Seleccção da semana: a semana seleccionada será a terceira semana do mês de Dezembro, por ser a semana representativa de toda a época alta não sendo a semana onde se espera ter

maior fluxo do mês, pois a maior frequência, em termos de fluxo estaria, em teoria, a semana de Natal.

2- PLANO DE AMOSTRAGEM

O plano de sondagem exige, geralmente, dois elementos de interesse:

- ✓ A determinação do tamanho da amostra
- ✓ A determinação da sua estrutura
 - O método de tiragem ou selecção dos elementos da amostra
 - A determinação da expressão do estimador do principal parâmetro θ (*téta*) que nos interessa estimar.

O método de tiragem é o procedimento escolhido para seleccionar os elementos que pertencerão a amostra. Um dos dados (elementos) fundamentais a ter em conta na escolha do método de tiragem dos elementos da população, são os valores das probabilidades de selecção ou de inclusão de cada elemento na amostra. Estas probabilidades serão posteriormente utilizadas na expressão do estimador para calcular as estimativas dos parâmetros desejados.

A expressão do estimador é a “fórmula” matemática escolhida em função dos elementos da amostra. Para um determinado método de tiragem existem numerosas expressões do estimador desejado; e uma expressão do estimador pode ser aplicado às diferentes amostras extraídas segundo métodos diferentes. A escolha do método de tiragem e da fórmula para o estimador dependem, geralmente, das considerações seguintes:

- ✓ Do enviesamento (precisão que se deseja para as estimativas);
- ✓ Da variância (variabilidade dos valores do estimador);
- ✓ Do custo associado à recolha dos dados;
- ✓ Da disponibilidade de informações auxiliares adequadas e actualizadas.

2.1-Tamanho e estrutura da amostra

A metodologia geral de determinação de um plano de amostragem consiste em se fixar os objectivos e calcular o tamanho da amostra e a estrutura “óptima” no sentido garantir os objectivos preconizados tanto em termos da precisão desejada, da representatividade para os domínios de estudo previamente estabelecidos, como no sentido de minimizar o custo total da pesquisa. O termo **representatividade duma amostra** é geralmente expresso, em termos estatístico, pelo **coeficiente de variação (C.V)**. Deste ponto de vista parte-se, geralmente, de uma estrutura de referência atribuída à uma plano de amostragem que segue os requisitos da amostragem aleatória simples (A.A.S) e, procura-se uma estrutura que se espera ser compatível com a pesquisa que se pretende realizar sem que se perca muito em termos de eficiência do plano amostral em relação à uma A.A.S.

2.1.1 Dimensão da Amostra: Aspectos gerais

Na determinação do tamanho da amostra, utiliza-se, geralmente, o **coeficiente de variação (C.V)** medido em termos de erro ou precisão relativa. O valor do C.V é frequentemente expresso e determinado em função do nível de confiança desejado e fixado à *priori*, do próprio parâmetro pesquisado e da própria variância deste parâmetro θ (*téta*), utilizando a seguinte formula ou transformação:

$$C. V_{\alpha} = \frac{z_{\alpha} \times \sqrt{V(\hat{\theta})}}{\theta} \quad (1)$$

Em que:

Z_{α} corresponde ao nível de confiança que se deseja e, é geralmente escolhido ou fixado: por exemplo, se nível de confiança é de 95%, então $Z_{\alpha} = 1,96$ se o nível de confiança é de 90% então $Z_{\alpha} = 1,645$. θ é o parâmetro estimado (um dos principais objectivos da pesquisa) na população em estudo, por exemplo média, a percentagem ou proporção, o total etc., e $V(\hat{\theta})$ é a variância estimada fornecida pela amostra. Um dos problemas bicudos dos estatísticos ou sondagistas e que constitui toda a questão da determinação do tamanho da amostra, reside exactamente no cálculo desta variância, pois ela depende exactamente do tamanho da amostra (n) da estrutura da amostra e da variabilidade da principal variável considerada (aquela que fornece os dados amostrais para estimar o parâmetro correspondente na população em estudo). Sendo assim, se não existirem

informações *à priori* suficientes e de qualidade que permitam determinar esta variância, o C.V ficará sem ser determinado e, conseqüentemente o tamanho da amostra poderá ficar condicionado.

Para remediar esta situação a dita variância é, geralmente, majorada no caso do principal parâmetro a estimar ser uma proporção (p) de valor igual a 0,25 (quando esta proporção é estimada ou fixada em 0,5). Pois, $V(\hat{\theta}) = P \cdot Q$ em que $Q = 1 - P$ ou seja que também é igual a 0,5. Por conseguinte, $V(\hat{\theta}) = P \cdot Q = 0,5 \cdot 0,5 = 0,25$. Nesta situação, C.V é geralmente expresso em termos de erro relativo e pode ser determinado por:

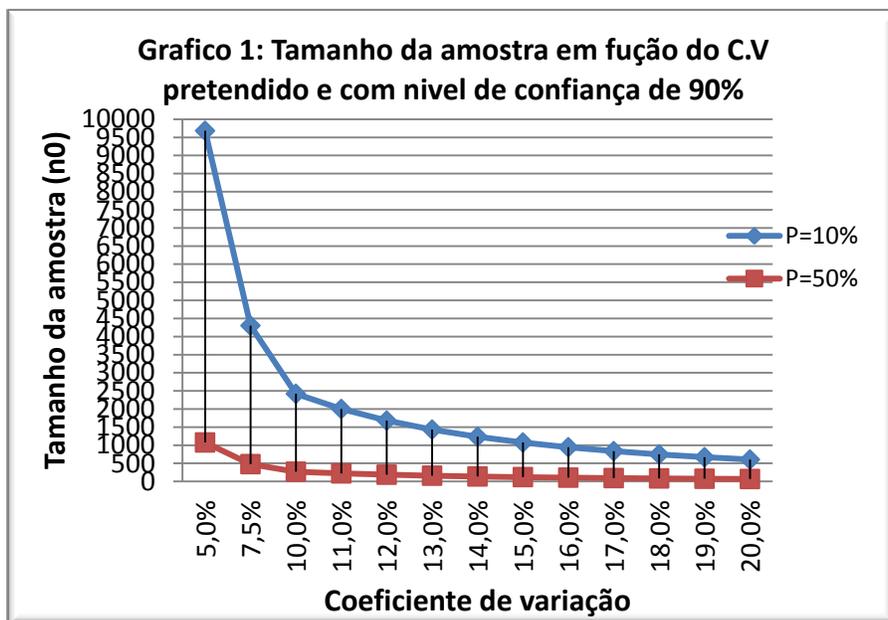
$$CV_{\alpha}^2 = Z_{\alpha}^2 \left[\frac{Q}{P \times n} - \frac{Q}{P \times N} \right] \quad (2).$$

(n) de uma amostra que se espera ser representativa, e que forneça dados para a estimativa dos principais parâmetros de interesse com o nível de confiança desejado e a precisão estabelecida,

através das fórmulas seguintes: $n_0 = \frac{Z_{\alpha}^2}{CV_{\alpha}^2} \times \frac{1 - P}{P}$ (3) em que $n = \frac{n_0}{1 + \frac{n_0}{N}}$ (4) servirá para

corrigir o tamanho da amostra quando a população estudada é finita(com N a designar-se o tamanho da população).

O gráfico 1 servirá como exemplos de diferentes tamanhos de amostra determinados quando se pretende estimar um parâmetro θ com um C.V desejado e nível de confiança 90%. Posteriormente, estes tamanhos poderão ser avaliados segundo os critérios (precisão, nível de confiança), corrigidos ou não, consoante se população estudada for considerada finita ou infinita, em termos da sua dimensão. Nota-se que se a população é muito grande (geralmente considerada superior a 100 000 indivíduos, a correcção não altera praticamente a dimensão inicial (n_0) ou seja a dimensão final n será aproximadamente igual a n_0): $n \approx n_0$. Por outras palavras, a correcção para a população finita é ainda mais importante, vantajosa e necessária quando menor for a população estudada. Após a correcção pela população finita confronta-se a dimensão da amostra encontrada com o orçamento disponível para inquirir (recolher dados) os n indivíduos (pessoas ou objectos).



Pode-se verificar pelo gráfico que quanto mais exigentes formos em termos de precisão (menor C.V) o tamanho da amostra, em S.A.S, terá que aumentar quase de maneira linear. Sendo assim se se pretender diminuir o C.V de 20% para 10%, por exemplo, teremos que aumentar o tamanho da amostra de cerca de 600 a 2420 indivíduos para se estimar uma percentagem na ordem dos 10%.

Expomos assim, os procedimentos básicos para a determinação do tamanho da amostra para uma pesquisa caso existam as informações de base necessárias e que se possa estimar ou fixar a variância do estimador ou quando se recorre à experiência de um *expert* (consultor) para a sua determinação ou estimação da forma como descrito acima. Estas são as principais razões do por que o tamanho da amostra ser muitas vezes fixado ao invés de ser determinado. A determinação da amostra é ainda mais complexa se se suspeitar da existência de um efeito de sondagem (denominado efeito de estrutura) que decorre basicamente em planos de amostragem complexa, por exemplo amostragem em várias etapas ou estágios comparativamente ao A.S.A. Este efeito que se denomina pela sigla DEFF (“design effect”) é frequentemente observado ou calculado quando “por natureza os elementos se agrupam para formar cachos ou conglomerados” proporcionando assim uma certa homogeneidade dentro de mesmo grupo. Muitas vezes há que se multiplicar a amostra inicial por um certo valor de DEFF para “restabelecer” alguma heterogeneidade desejável em planos de sondagem e garantir assim um pouco de independência e aleatoriedade dos elementos na amostra. O que faz com que a estrutura da amostra no plano de amostragem utilizado seja “plano

pseudónimo” dum plano que respeite os requisitos da Sondagem Aleatória Simples (S.A.S) em que o seu DEFF é a referencia e igual à unidade.

Para este inquérito concreto a dimensão da amostra não é determina usando a fórmula usual supra descrita porque 1) não se conhece o P (proporção de uma variável de interesse) que poderia ser por exemplo a percentagem de turistas que declararam os gastos, ou a percentagem de turistas com gasto superior a um determinado valor, e 2) nem se pretendia fixa-lo à *priori* a, por exemplo 50% (apesar de ser recomenda nas bibliografias quando não se tem uma estimativa actual e fiável para o valor do *p*), porque a população em estudo (turistas de 15 anos ou mais que visitam o país) é relativamente variável e depende em larga escala da conjuntura nacional e internacional. Outrossim e particularmente por esta ultima razão o tamanho da amostra para o ISGT não foi fixado à *priori* em termos absolutos, por exemplo em 1000, 1500 ou 2000 entrevistas para depois se verificar o erro relativo (C.V) na estimativa do parâmetro de interesse com um dado nível de confiança.

No entanto, recorreu-se a experiencia de um *expert* que nos ajudou a fixar o tamanho da amostra inicial (*n*₀) em termos de percentagem correspondendo a taxa de sondagem constante de 10% do total dos passageiros dos voos internacionais que depois serão adicionados para se obter o total de entrevistas de todos os voos.

2.1.2- Estimação do tamanho da amostra inicial (*n*₀)

Dado a série de dados que indica o total de passageiros embarcados no ano anterior (por ser o mais actualizado possível) por cada voo e tipo de voo (charter ou regular em cada um dos 4 domínios de estudo) calcula-se a taxa de crescimento de passageiros segundo estas categorias através de uma média móvel até o mês de Setembro do ano 2010 e de seguida calcula-se a média dessas médias móvel. O valor da média será posteriormente utilizado como factor para prever a dimensão da população (*N*) que representará o número de passageiros embarcados previsto (ou esperado) para o mês homólogo em 2011 por cada voo e tipo de voo (charter ou regular em cada um dos 4 domínios de estudo) segundo a fórmula seguinte.

$$T_{t,2011} = r * T_{t,2010} \quad (5)$$

Em que:

$T_{(t,2011)}$: é o total previsto de passageiros embarcados de cada voo e tipo de voo em cada um dos 4 domínios de estudo para o mês de Setembro tendo em conta o número de passageiros embarcados observados em Setembro de 2010.

$T_{(t,2010)}$: é o total de passageiros embarcados observados de cada voo e tipo de voo em cada um dos 4 domínios de estudo em Setembro de 2010 segundo as mesmas categorias.

r : representa o valor, o rácio entre o índice médio de variação do fluxo de turistas calculados para os n meses anteriores ao mês homólogo (Setembro) de 2010 segundo as mesmas categorias e o índice médio de variação do fluxo de turistas calculados para os n meses de 2011 anteriores ao mês em que se realiza o estudo (Setembro) segundo as mesmas categorias ou seja:

$$r = \frac{I_{\text{jan,agost em 2010}}}{I_{\text{jan,agost em 2011}}} \quad (6)$$

Com o valor $T_{(t,2011)}$ estimado calcula-se o valor inicial da amostra (n_0) de passageiros embarcados de cada voo e tipo de voo em cada um dos 4 domínios de estudo para o mês de Setembro, segundo a formula seguinte.

$$n_0 = 0,1 * T_{t,2010} \quad (7)$$

Uma vez obtido o plano de voo com a informação sobre a origem e a capacidade do voo deve-se proceder, por domínio de estudo e tipo de voo da seguinte forma relativamente à taxa de sondagem (ou resposta efectiva):

Estratificação da amostra: serão considerados dois estratos segundo o tipo de voo.

- 1) Voo charter
- 2) Voo regular

Repartição da amostra por estrato

Quadro 1: repartição da amostra por domínio de estudo e estratos

Estratos	Domínios de estudo		
	Sal	Boa Vista	Praia
Charter	8%	8%	10%
Regular	12%	12%	10%
Total	10%	10%	10%

O tamanho n inicial que se espera inquirir poderá ser sobreavaliados pois o cálculo incidu sobre um total que contempla também as crianças de menos de 15 anos (recorda-se que as entrevistas validas são para as pessoas não residentes de 15 anos ou mais). Mas também poderá ser sobreavaliados pois baseou-se numa “projectão” do número de turistas por voo e tipo de voo em cada um dos 4 domínios de estudo com base em observações passadas e sob hipótese de se manter uma certa regularidade quanto ao número de passageiros captado pelo valor de r (ver equação 6). No entanto, após as entrevistas pode-se ajustar e ter um n final superior ao n_0 inicial. Dado que o objectivo principal deste inquérito é a estimação de gasto realizados por turistas não residentes, há necessidade de se corrigir ou se extrapolar os dados do inquerido por todos os potenciais turistas não residentes que visitaram o país em 2011.

2.1.2.1 Correção da amostra inicial (n_0) para a população finita: obtenção de amostra inicial corrigida (n)

Uma vez obtida uma estimação da amostra de turistas através da fórmula (7) pode-se utilizar a

seguinte formula: $n = \frac{n_0}{1 + \frac{n_0}{N}}$ para obter o tamanho da amostra inicial de turistas corrigida que se

vai inquirir. Infelizmente, como só se conhece o N (total de turistas efectivamente embarcados), por voo, por aeroporto, por mês e por tipo de voo, depois de realizarmos o inquérito, não se procede a correcção da fórmula (7) pela fórmula de n proposta. Sendo assim, espera-se inquirir o n_0 turistas como previamente estimado. No entanto, como a amostra final de turistas que será utilizada para estimar os parâmetros de interesse, por exemplo o gasto médio de turistas, faz referência

exclusivamente aos **turistas** não residentes de 15 anos ou mais, ela será corrigida pela estrutura (em percentagem) de turistas não residentes de 15 anos ou mais efectivamente intuídos na amostra n_0 (este ponto será retomado quando se aborda o processo de estimação dos parâmetros).

2.1.2.2-Estrutura da amostra inicial

A estrutura de referência dum S.A.S não será a estrutura “real” da amostra da pesquisa ISGT, porque os indivíduos serão submetidos à uma estrutura de selecção mais ou menos complexa em termos de estágios (etapa) de selecção e da estratificação da amostra. A questão principal seria: deve-se escolher uma estrutura a duas etapas (1ª etapa: selecção dos voos, e 2ª selecção dos visitantes) ou à uma só etapa onde se selecciona directamente os turistas? A segunda opção seguia mais ou menos uma amostragem alegoria simples em que seleccionava-se os turistas independentemente dos voos. Não era preciso muitas informações adicionais, mas seria indispensável uma listagem actualizada dos passageiros a embarcar. Dado aos problemas que se prendem principalmente com a segurança dos passageiros e a protecção de confidencialidade dos dados individuais é muito difícil aceder à esta listagem onde se poderia identificar e seleccionar os turistas. Atendendo à imitação dum estrutura da amostra aleatória simples devido à inexistência dum listagem actualizada das unidades amostras a estrutura amostral sugerida para este inquérito assemelha-se à uma estrutura de uma sondagem em duas etapas com o particular interesse de na primeira etapa ser seleccionada exactamente todos as unidades primarias que são os voos e na segunda etapa seleccionar 10% do total de indivíduos em cada aeroporto segundo a repartição descrita no quadro 2 em baixo.

2.2- Base de Amostragem e População Alvo

Na verdade não é possível obter a base de amostragem (uma lista exhaustiva de todos os passageiros segundo algumas características que permitissem identificar cada um dos passageiros). Dispõe-se de um ficheiro com indicação do número total de passageiros embarcados (residentes e não residentes) construído a partir de informação fornecida pela ASA/TACV. Este ficheiro é composto por um conjunto de campos que permite a construção do seguinte quadro:

Quadro 2 - Informação da Fronteira Aérea (Aeroporto XXXX)

Data Mês XXXX	Aeroporto de destino	País de destino (voo)	Efectivo (Embarcados)	Tipo de voo(1=charter ; 2 regular)

Dispõem-se assim de um conjunto de dados indicando o total de população (total de indivíduos viajados por voos e tipo de voo, aeroporto de destino, país de destino e mês).

Características da amostragem:

- Tipo de amostragem:

Para este inquérito considerou-se uma **amostragem probabilística** pois, embora a probabilidade de selecção de um elemento da amostra não seja conhecida à *priori*, pode-se fixa-la. Efectivamente esta probabilidade é fixada em 10%. Os argumentos principais são: os voos (conglomerados) são em teoria homogéneos e quando se aplica uma percentagem inferior poderia correr o risco de estar a seleccionar pessoas com características muito semelhantes uma vez que as pessoas que viajam com amigos, geralmente, sentam-se juntas, têm características teoricamente parecidas, por exemplo é forte provável que sejam igualmente propensas ao gasto. E, se tivéssemos que seleccionar, na primeira etapa, N voos de um conjunto de M voos para depois seleccionar, na segunda etapa, os m indivíduos em cada voo seleccionado, teríamos $\bar{n} = \frac{1}{M} \sum_{i=1}^M M_i$, (o número médio de passageiros nos voos seleccionados e M_i seria o número de passageiros do voo i) e esta opção podia levar à situação de se inquirir muitas pessoas por voos. Isto pode não ser viável principalmente porque, em teoria, os voos são relativamente homogéneos no seu interior: por exemplo os voos só com franceses ou só com espanhóis, ou ingleses etc., e, diferentes (heterogéneos) de um voo para outro. Teoricamente, entre os franceses, entre os ingleses, ou entre os espanhóis, por exemplo, o gasto médio pode ser parecido (relativamente próximo), mas pode ter diferença relativamente importante, por exemplo, em relação do gasto médio de um francês comparativamente a um inglês ou espanhol.

Por exemplo, se se pretende inquirir 1000 visitantes no total, poder-se-ia selecciona-los aleatoriamente (solução mais eficaz mas pode ser mais custoso). Seleccionar na primeira

etapa 500 voos e depois 4 visitantes por voo deveria deteriorar um pouco a qualidade dos resultados, extrair 250 voos e depois inquirir 4 visitantes deveria ser pior que a solução anterior sendo que a pior solução seria seleccionar um único voo e inquirir as 1000 pessoas (este caso não seria também possível pois não se tem voos com 1000 pessoas). Sendo assim, as literaturas de sondagens tenta estabelecer a seguinte regra: para o tamanho de amostra global fixo em que $n_0 = m * \bar{n}$ preferível escolher o máximo m e o \bar{n} mínimo mas com reflexo directo sobre o custo (ARDILLY,1994:99).

No caso do ISGT optou-se por uma solução que vai no mesmo sentido de seleccionar todas as unidades primárias (UP) ou todos os voos (m seria máximo!) com intuito de assegurar uma taxa de sondagem “considerada razoável” e constante em todos os voos: de 10% do total de passageiros, em cada aeroporto assegurando-se assim o \bar{n} mínimo.

3-ALGUNS ASPECTOS OPERACIONAIS DO INQUÉRITO

3.1 Selecção dos indivíduos: amostra inicial

O método de tiragem ou selecção de visitantes em cada voo, por tipo de voo e em cada domínio de estudo será uma selecção **sistemática simples** em que cada indivíduo será seleccionado respeitando um passo de sondagem de 10%, ou seja, selecciona-se 1 em cada 10 turistas de 15 anos ou mais completos (residentes e não residentes) que saem do país, nos voos internacionais. A selecção de turistas é feita no momento da entrevista, durante uma semana (ou seja sete dias, de forma a obter informação para todos os dias da semana).

Como não se dispõem de uma listagem de turistas identificando-lhes minimamente (nome, sexo, nacionalidade, etc,) de onde se enumera de 1 a N os indivíduos e rotulá-los com um número de identificação, o processo de selecção sistemática simples utilizada neste inquérito não respeitara todas as técnicas formais, por exemplo a selecção de um número aleatório entre 1 a K, sendo K o passo de sondagem.

Neste caso, o procedimento de selecção proposto é o seguinte:

- a) Previamente à selecção dos visitantes e a respectiva entrevista, prodem-se que o chefe de equipa obtenha junto dos serviços de *chek-in*, o plano de voo (qual é o voo previsto para as próximas horas indicando também o total previsto (N) de passageiros a embarcar.
- b) Cada equipa de recolha (inquiridores e chefe de equipa) deve ir se fixar à saída da porta de “Raio X” ou passagem magnética.

Como já se sabe que o passo de sondagem é de 1 em cada 10 visitantes que passam pela porta do “Raio-X” (ou seja $K=10$) e este passo de sondagem é constante de cada vez que se vai seleccionar os visitantes. Em princípio os visitantes começam a dirigir-se à sala de embarque uma hora antes da partida para o avião (descolagem) de maneira aleatória e, teoricamente, segundo uma distribuição de probabilidade de Poisson (λ): em que λ pode significar o número médio de turistas que passam ao “Raio-X” por intervalo de tempo, independentemente de existirem ou não sobreposição de voos despachados.

Obs 1): Deve-se garantir (existir) o número mínimo de 30 passageiros por voos. Consequentemente os voos com menos de 30 passageiros não fazem parte do âmbito de inquérito.

Obs 2): No caso de existir voos com 30 a 59 pessoas, ao seleccionar 10% seria equivalente seleccionar de 3 a 5 pessoas. Nestes casos deve-se manter o mínimo de 5 pessoas na amostra, por isso deve-se seleccionar 5 pessoas. Estas 5 pessoas devem ser seleccionadas com algum espaçamento temporal. Por exemplo, se um voo estiver previsto para sair às 10h A.M com uma duração de duas horas antes para se dirigir à sala de embarque terá que distribuir as entrevista proporcional ao tempo de espera antes da partida para o avião (descolagem). Por exemplo, se o período destinado para se dirigirem à sala de embarque for de 60min deve-se dividir 60min por 5. O que significa que se deve seleccionar 1 pessoa em cada 12 minutos. Esta regra não é rígida mas com auxílio dos chefes de equipas e supervisores vai se encontrar o melhor espaçamento e repartição de pessoas a seleccionar por período segundo o número de passageiros que se dirigem à sala de embarque.

3.2- Como seleccionar um turista?

3.2.1 Em voos despachados isoladamente (um voo em cada horário)

Quer dizer, depois de estarmos apostos para a selecção (questionários preparados, ter chegado à passagem de “Raio-X” como é que estas 5 pessoas deverão ser seleccionadas? Devemos Selecciona-se a 1ª, 11ª; 21ª; 31ª; 41ª que passar na porta do “Raio-X” (totalizando as 5 pessoas que devíamos seleccionar e inquirir)? Devemos seleccionar a 2ª, 12ª; 22ª; 32ª; 42ª que passar na porta do “Raio-X” (totalizando as 5 pessoas que devíamos seleccionar e inquirir)? Ou outra sequência respeitante um passo de 10! Por exemplo a 10ª; 20ª ; 30ª; 40ª ; 50ª que passar na porta do “Raio-X” (totalizando as 5 pessoas que devíamos seleccionar e inquirir)? Ou ainda uma outra sequencia de 10!.

Naturalmente podemos seleccionar directamente as primeiras 5 pessoas. Mas devemos respeitar o passo de sondagem “seleccionar 1 em cada 10 turistas”, como justificado previamente. Neste caso, como não devemos concentrar as entrevistas num único período, pretende-se que a selecção seja também representativa ao longo do tempo destinado à entrevista (põe exemplo uma hora antes da chamada para o avião). Neste caso os intervalos de tempos iguais servem também como critério para, em teoria, conferir aleatoriedade e representatividade das entrevistas ao longo do período consignado ao inquérito.

Portanto o passo de sondagem “seleccionar 1 em cada 10 turistas”, deve ser mantido dentro dos períodos de tempos iguais (12 em 12 minutos). No nosso exemplo de referência, devemos seleccionar 1 turista em cada 12 minutos. Será que se pode seleccionar logo no decorrer do primeiro minuto? A resposta é afirmativa. Uma vez que tudo esteja apostado para a selecção e entrevista, **esta pessoa pode ser seleccionada a qualquer momento nos primeiros 12 minutos**. A segunda pessoa nos próximos 12 minutos, a qualquer momento, e assim por diante. Neste caso, como por hipótese os turistas passam ao Raio-X de maneira aleatória e os inquiridores não tem nenhuma informação prévia quanto às características de turistas a seleccionar (a não ser que a condição necessária em que os turistas deva ter 15 anos ou mais), a selecção da primeira, da segunda e da **j-ésima** pessoa, sendo uma em cada período de 12 minutos, ficará ao critério do inquiridor sob a coordenação de cada chefe ou supervisor de cada equipa. Naturalmente, como na pratica nem sempre esta regularidade é observada, dependendo da experiencia da perspicácia dos inquiridores

e do chefe da equipa eles terão alguma margem para proceder algumas alterações metodológicas que se impuserem.

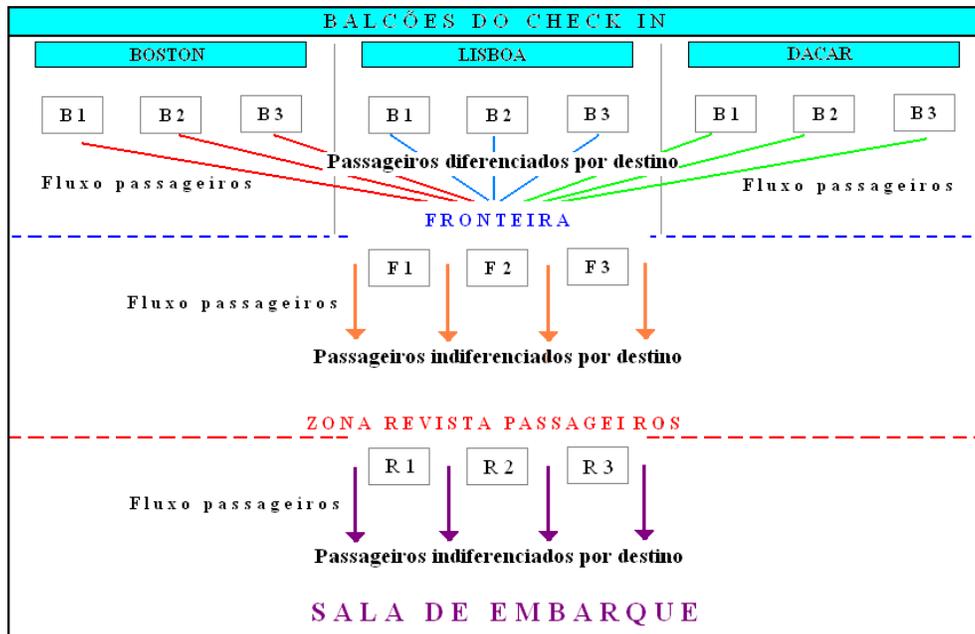
Obs 3): Se os turistas vão sendo seleccionados dos voos com 60 a 100 pessoas, então 10% seria equivalente a seleccionar 6 a 10 pessoas. Neste caso deve-se seleccionar 10 pessoas. As mesmas regras exemplificadas acima, devem aplicar. Neste caso deve-se seleccionar 1 pessoa em cada 10 minutos se formos seleccionar turistas de um voo de 60 passageiros, ou 1 em cada 6 minutos se forem voos de 100 pessoas. Para os voos com mais de 100 pessoas deve-se seleccionar os 10% como previsto neste plano amostral. Os mesmos procedimentos concernentes à selecção das pessoas e a repartição das entrevistas por período de tempo, devem ser aplicados segundo cada caso.

O procedimento respeitando ao espaçamento no tempo (períodos antes do embarque e os dias de semana) são para serem mantidas pois o objectivo principal não é terminar as entrevistas o mais rapidamente possível mas sim, ter a representatividade de inquéritos durante todo o período de afluência à sala de embarque, durante todo o dia em que haja voos e durante todos os 7 dias de semana. A selecção dos turistas terá sempre que respeitar a taxa de sondagem: seja 1 em cada 10 pessoas e sempre à passagem de “Raio-X”.

3.2.2 Em voos despachados conjuntamente (um voo em cada horário)

Em princípio, caso não haja sobreposição de voos despachados a afluência dos turistas à sala de embarque se faça de maneira, organizada, aleatória consoante o tipo de voo, destino e horário de partida. A figura 1 mostra o fluxo de passageiros para este caso.

Figura: 1 Esquema actual de movimento do passageiro desde balcão de Check in, até à sala de embarque, nos momentos em que há mais de um voo a ser despachado em simultâneo.



Obs 1: Nesta situação a proposta de selecção dos turistas obedece os mesmos princípios gerais descritos acima para caso de voos despachados isoladamente. A grande diferença é que os turistas terão que ser seleccionados segundo o destino do voo. Por exemplo, se tivermos que seleccionar 20 (10%) pessoas num voo para Boston em 1 uma hora, devemos seleccionar 1 pessoa em cada 3 minutos.

Obs 2: Como os turistas passam à porta de "Raio-X", indiferentemente segundo o destino de voo, há que certificar-se dos que vão para Boston (perguntando-se e /ou verificando-se o bilhete de embarque). Neste caso, ainda, como a passagem é aleatória pode ser que nos primeiros 3 minutos por exemplo, não passe turistas com destino ao Boston. Por conseguinte, ficaria difícil respeitar os procedimentos metodológicos como descrito (de seleccionar 1 turista em cada 3 minutos). Por isso, é provável que em caso de voos simultâneos tenhamos necessidade de concentrar entrevistas em períodos específicos de forma a realizar com êxito todas as selecções e entrevistas. Mas sempre que possível estes procedimentos metodológicos devem ser mantidos em prol da eficácia do plano de amostragem.

3.3 Observações importantes: resumo

Ao longo da semana de referência, os entrevistadores deverão realizar um número específico de entrevistas (sem qualquer restrição relativamente ao passageiro, nomeadamente no que concerne ao país de residência do mesmo) nos vários voos especificados. As entrevistas devem ser “expandidas” de forma a cobrir todas as horas do voo e em todos os dias da semana. Quer dizer, não é aconselhável que se concentrem as entrevistas num único dia ou em alguns dias e num único período ou em algumas horas. Em cada aeroporto, as pessoas serão inquiridas independentemente dos voos são e do número de passageiros a embarcar desde que o voo tenha no mínimo 30 passageiros.

Se o voo partir de mais do que um aeroporto internacional (por exemplo, se sair da praia e fizer escala no Aeroporto do Sal) com passagem pelos serviços de fronteira, o inquérito será realizado no aeroporto onde o passageiro passa pela última vez pelos serviços de fronteiras.

Note-se que apesar de a amostra ser estratificada apenas por tipo de voos (charter ou regular), todos os voos serão observados independentemente do tipo de voo.

- **Insistências/Tratamento de recusas:** não aplicável.
- **Possibilidade de subsituação ou de inquirir outro turista por indicação do turista seleccionado:** não admitido.
- **Utilização de incentivos:** não aplicável.
- Se o passageiro que estiver a preencher o questionário for chamado a embarcar deve-se considerar o questionário fechado.
- Em caso de adiamento dum voo despachado e com o inquérito já realizado, o inquérito é válido e não será mais realizado no voo de substituição.
- Caso haja permuta de voos, observa-se na mesma o voo que vai acontecer no período ou horário estabelecido.

4- DETERMINAÇÃO DA AMOSTRA FINAL

A amostra final será o somatório das entrevistas válidas recolhidas junto de várias sub-amostras de proporções iguais e, fixado todas a 10% do total de passageiros de cada voo registado em cada um dos três domínios de estudo segundo os dois estratos escolhidos.

4.1-Amostra final de cada domínio

$$n_{\text{Final,Domineo}} = n_{\text{Final,Charter}} + n_{\text{Final,Regular}} \quad (8)$$

E

4.2-Amostra final Global

$$n_{\text{Final}} = n_{\text{Final,Sal}} + n_{\text{Final,BoaVista}} + n_{\text{Final,Svicente}} + n_{\text{Final,Praia}} \quad (9)$$

4.3 Ponderação à *posteriori* das proporções de turistas não residentes

A amostra inicial de turistas contempla todos os turistas de 15 anos ou mais completos sejam eles residentes ou não residentes. A inferência estatística sobre um dos principais parâmetros de interesse, por exemplo o gasto médio é feita ao universo dos turistas de 15 anos ou mais completos que sejam não residentes. Por isso há a necessidade de se ponderar as entrevistas realizadas de forma a reflectirem a estrutura dos turistas não residentes.

Após realização do inquérito conhece-se, na amostra final inquirida, a verdadeira proporção de turistas de 15 anos ou mais completos residentes e dos não residentes, assim como outras características. Conhece-se também a verdadeira população de turistas embarcados por voo, destino do voo, mês do voo (N). Dado que amostra é esperada ser “representativa”, pode-se fazer uma correcção da repartição à *posteriori* das proporções de turistas não residentes através de um factor de correcção.

4.3.1-Cálculo do factor de ponderação ou de expansão

Com os dados da ASA relativo às estatísticas de voos internacionais procura-se preferencialmente encontrar o número total de pessoas embarcados por voos (charter e regular), por aeroporto, ou na pior das hipóteses o total por mês. Se for possível encontrar estes dados, o factor de extrapolação poderá ser calculado da seguinte forma.

$$f_j = \frac{W_j * T_j}{n_j}$$

Em que,

f_j : é o factor de extrapolação em que se multiplicara por cada registo correspondentes aos indivíduos de 15 anos ou mais na amostra.

W_j : é o peso ou a percentagem dos turistas não residentes inquiridos na amostra segundo a época j ($j=1$: época baixa, $j=2$: época alta)

T_j : é o total de pessoas efectivamente embarcadas (observados) por voo fornecido pela ASA para o mês e ano de estudo segundo o nível de detalhe supracitado.

n_j : é o total de entrevistas validas (incluindo os não residentes).

ANEXO A: OUTROS CONCEITOS IMPORTANTES

Excursionista

Visitante que não pernoita no lugar visitado.

Notas: Inclui os passageiros em cruzeiro que permanecem em navios, bem como os membros das respectivas tripulações.

Motivo Principal da Viagem

Motivo na ausência do qual a viagem não se teria realizado. São contemplados os seguintes motivos:

- Lazer, Recreio e Férias: repouso, gastronomia, compras, desporto como espectador e prática de desporto, educação, encontros (não profissionais), cultura e entretenimento como espectador, artes, hobbies, jogos e outros (não profissionais);
- Profissionais/Negócios: reuniões, convenções, seminários, conferências, congressos, feiras e exposições (participação profissional), missões, viagens de incentivo, vendas, marketing e outros serviços, pesquisa, ensino, consultoria, cursos de idiomas, educação, investigação, profissionais artísticos, culturais, religiosos e desportivos;
- Visita a Familiares e Amigos: visitas a familiares e/ou amigos, participação em funerais, casamentos, aniversários e outros eventos familiares;
- Saúde (razões voluntárias): tratamentos e cuidados de saúde em estâncias termais, balneares, lares de convalescência e outros tratamentos e curas;
- Religião e Peregrinação (não profissional): assistência a eventos religiosos e peregrinação;
- Outros Motivos.

Nacionalidade

Cidadania legal do indivíduo no momento de observação; são consideradas as nacionalidades constantes no passaporte, na autorização de residência ou no certificado de nacionalidade apresentado. Os indivíduos que, no momento de observação, estejam com um processo de naturalização em curso devem ser considerados com a nacionalidade que detinham anteriormente.

País de Residência

Uma pessoa é considerada residente de um país (local) se:

- a) tiver vivido pelo menos 6 meses nesse país (local), ou
- b) tiver vivido nesse país (local) por um período mais curto mas que pretenda regressar no prazo de 6 meses, com a intenção de se instalar nesse país/local.

Permanência Média

Número de noites que os turistas permanecem, em média, numa região ou país.

Saída de Residentes

Número de indivíduos residentes em Cabo Verde que saem do país num dado período, seja qual for o motivo da viagem.

Turismo

Actividades realizadas por indivíduos durante as suas viagens e estadas em lugares distintos da sua residência habitual, por um período de tempo consecutivo inferior a um ano com fins de lazer, negócios ou outros motivos.

Turismo Emissor

Inclui as actividades dos residentes de um determinado país noutros países, fora do seu ambiente habitual.

Notas: Este conceito aplica-se igualmente a uma região.

Turismo Internacional

Inclui as actividades dos visitantes residentes no estrangeiro que viajam num outro país, fora do seu ambiente habitual (turismo receptor), e as actividades dos residentes de um determinado país noutros países, fora do seu ambiente habitual (turismo emissor).

Notas: Este conceito aplica-se igualmente a uma região.

Turismo Receptor

Inclui as actividades dos visitantes residentes no estrangeiro que viajam num outro país, fora do seu ambiente habitual.

Notas: Este conceito aplica-se igualmente a uma região.

Turista

Visitante que permanece pelo menos uma noite num alojamento colectivo ou particular no lugar visitado.

Viajante

Indivíduo que se desloca entre dois ou mais países distintos ou entre dois ou mais lugares no interior do seu país de residência habitual, independentemente do seu motivo.

Visitante

Indivíduo que se desloca a um lugar diferente da sua residência habitual, por uma duração inferior a 365 dias, desde que o motivo principal da viagem não seja o de exercer uma actividade remunerada no lugar visitado.

Notas: O termo visitante inclui: turistas e excursionistas. Os três critérios fundamentais para distinguir os visitantes de outros viajantes são os seguintes:

- i) a deslocação deve efectuar-se a um local diferente do ambiente habitual do indivíduo;
- ii) a estada no local visitado não deve ultrapassar doze meses consecutivos;
- iii) o objectivo principal da visita não deve ser o exercício de uma actividade remunerada no local visitado.