



# **OBSERVATORIO TURÍSTICO DE BAJA CALIFORNIA: FASE III**

## **REPORTE**

**METODOLOGÍA Y DISEÑO MUESTRAL DE LA  
ENCUESTA DE VISITANTES NACIONALES EN  
LAS CIUDADES FRONTERIZAS DE BAJA  
CALIFORNIA: MEXICALI, TECATE Y  
TIJUANA  
(VERANO 2013)**





Observatorio Turístico de Baja California: Fase III

# **OBSERVATORIO TURÍSTICO DE BAJA CALIFORNIA: FASE III**

## **REPORTE**

**METODOLOGÍA Y DISEÑO MUESTRAL DE LA  
ENCUESTA DE VISITANTES NACIONALES EN  
LAS CIUDADES FRONTERIZAS DE BAJA  
CALIFORNIA: MEXICALI, TECATE Y TIJUANA  
(VERANO 2013)**

Tijuana, Baja California, 2013



## **Observatorio Turístico de Baja California: Fase III**

**Coordinadora general:**  
Dra. Nora L. Bringas Rábago

**Coordinador adjunto:**  
Dr. Djamel Toudert

**Participantes:**  
Dr. Basilio Verduzco Chávez  
Mtra. Yessica Vicencio Murillo  
Mtra. María Teresa López Avedoy  
Mtra. Mariana Trejo Sánchez  
Mtro. Christian Angeles Salinas  
Lic. Mariana Argüello Escobedo  
Lic. Alan Ingram Campillo

**Apoyo técnico especializado:**  
Lic. Carlos Vladimir Ruelas González  
Lic. Alberto Joaquín Acosta Rosales

**Coordinación de revisión y captura:**  
Lic. Luis Francisco Lares Serrano

**Marco muestral:**  
Mtro. Gilberto Hernández Olvera

**Coordinación de trabajo de campo:**  
Mtra. Aurora Guadalupe Villa Aparicio  
Mtro. Heber Huizar Contreras

**Diseño gráfico:**  
Lic. Claudia Ramos Castillo

**Proyecto financiado por la Secretaría de Turismo del Estado de Baja California**

**Secretario de Turismo:**  
C.P. Oscar Escobedo Carignan

**Subsecretario de Turismo:**  
Lic. Ives Lelevier Ramos



## Observatorio Turístico de Baja California: Fase III

### Índice

Introducción .....	1
1. Diseño conceptual .....	2
1.1 Objetivos .....	2
1.2 Población objetivo .....	3
1.3 Cobertura .....	3
1.4 Población no cubierta .....	3
1.5 Tipo de investigación .....	4
2. Diseño metodológico .....	4
2.1 Marco muestral .....	4
2.1.1 Estimación del flujo de pasajeros en centrales de autobuses .....	8
2.1.2 Estimación del flujo de pasajeros en terminales aéreas .....	10
2.2 Estratificación del espacio geográfico y cálculo de probabilidades de primer orden. ....	11
2.2.1 Centrales de autobuses .....	11
2.2.2 Terminales aéreas .....	13
2.3 Tamaño de muestra .....	15
2.3.1 Afijación de la muestra .....	17
2.3.2 Selección de la muestra .....	17
2.4 Mediciones .....	18
2.4.1 Proceso de estimación .....	19
2.4.2 Factor de discriminación .....	19
2.4.3 Factor geográfico .....	20
2.5 Factor temporal .....	21
Referencias .....	24

### Índice de cuadros

Cuadro 1. Marco muestral de la central de autobuses de la ciudad de Tijuana, B.C. ..	7
Cuadro 2. Marco muestral de la central de autobuses de la ciudad de Mexicali, B.C. ..	7
Cuadro 3. Marco muestral de la central de autobuses de la ciudad de Tecate, B.C. ....	8



## Observatorio Turístico de Baja California: Fase III

### Introducción

El Colegio de la Frontera Norte y la Secretaría de Turismo del Estado de Baja California desarrollaron en convenio el proyecto denominado “*Observatorio Turístico de Baja California*”, encaminado a generar, sistematizar y presentar información relevante, confiable y de calidad sobre las características del turismo nacional en las ciudades de Mexicali, Tecate y Tijuana, Baja California.

En este contexto, el siguiente documento tiene como objetivo mostrar la metodología y el diseño muestral de la encuesta “Caracterización de los visitantes nacionales en las ciudades fronterizas de Baja California: Mexicali, Tecate y Tijuana (verano 2013)”<sup>1</sup>. Para dar cumplimiento a este objetivo se realizó un estudio detallado de la dinámica de movilidad de pasajeros en cada punto de salida, como puertos aéreos y terrestres. Este análisis permitió establecer el diseño muestral de la encuesta de visitantes con las siguientes características:

- Captura de la información del viaje de la población objeto de estudio que se encontraba a punto de partir hacia un destino diferente al punto de muestreo.
- Diseño de un estudio estrictamente probabilístico que recogiera la variación de manera individual de cada uno de los puntos de muestreo, el cual permitiera generar inferencias particulares de cada estrato o dominio de estudio caracterizado.
- Utilización de un esquema de muestreo probabilístico, estratificado y bietápico.
- En cada punto de muestreo se aplicó una técnica de abordaje con las siguientes características:

---

<sup>1</sup> Esta encuesta tuvo como antecedente el estudio “Caracterización de los visitantes nacionales en las ciudades fronterizas de Baja California: Mexicali y Tijuana (invierno 2012-2013)”, Reporte de investigación, El Colef- Secture, del cual se retomaron los puntos y zonas de muestreo para las ciudades de Tijuana y Mexicali, agregando para este estudio a la ciudad de Tecate.



## **Observatorio Turístico de Baja California: Fase III**

- Centrales de autobuses: las encuestas se aplicaron en salas de espera mediante una selección sistemática del entrevistado.
- Aeropuertos: las encuestas se aplicaron tanto en la puerta de entrada, como en la sala de espera posterior a los controles de revisión y seguridad del aeropuerto.
- La recolección de la información se realizó de manera mensual y trimestral para captar el comportamiento de las diferentes estaciones del año.
- La información fue sometida a estrictos controles de calidad durante el operativo de campo, como en el proceso de estimación.
- El proceso de estimación siguió las mismas pautas establecidas en el diseño muestral, como en la selección de la muestra.

Los aspectos detallados sobre el marco conceptual y la metodología de la encuesta Caracterización de los visitantes nacionales en las ciudades fronterizas de Baja California: Mexicali, Tecate y Tijuana (verano 2013) se presentan a continuación.

### **1. Diseño conceptual**

#### **1.1 Objetivos**

- Dimensionar el número de turistas nacionales que llegaron a la región fronteriza de Tijuana, Mexicali y Tecate durante los meses de julio, agosto y septiembre de 2013.
- Determinar las características sociodemográficas y económicas de los visitantes nacionales en la región fronteriza de Mexicali, Tecate y Tijuana.
- Determinar tanto el uso como la valoración de los servicios turísticos y públicos utilizados por los turistas nacionales en las ciudades de estudio.
- Cuantificar el gasto promedio de los turistas nacionales en Baja California.
- Determinar las características de estancia como la valoración del uso de servicios públicos y turísticos.



## **Observatorio Turístico de Baja California: Fase III**

### **1.2 Población objetivo**

Personas de 15 o más años que viven en México, con residencia habitual en una ciudad distinta a la de la entrevista, y que después de visitar Tijuana, Mexicali o Tecate, estaban retornando a su lugar de origen desde las centrales de autobuses y terminales aéreas; que no trabajaron o desempeñaron alguna actividad remunerada en las ciudades de estudio y que utilizaron servicios considerados como turísticos.

### **1.3 Cobertura**

El estudio tuvo como alcance geográfico las centrales de autobuses y terminales áreas de las ciudades de estudio. Dependiendo de la cobertura espacial, la encuesta se clasifica en dos modalidades que son: turistas en centrales de autobuses y turistas en terminales áreas. Ambas poblaciones forman parte de la encuesta nacional. La encuesta tuvo como alcance temporal el periodo que comprende del 1º de julio al 30 de septiembre de 2013. Los niveles para el que son válidas las inferencias, corresponden a la localidad de muestreo, mes y modalidad de turista.

Los horarios de aplicación de la encuesta cubren los flujos de personas entre las 00:01 a.m. y 23:59 p.m., es decir, todo el día. Derivado de este efecto temporal, al final se tuvo una matriz de horarios de muestreo probables de 830 filas por 20 columnas en promedio para las localidades estudiadas.

### **1.4 Población no cubierta**

En todas las ciudades se cubrió completamente el espacio geográfico, el cual estuvo determinado por el flujo de pasajeros en las centrales de autobuses y en las terminales aéreas.



## **Observatorio Turístico de Baja California: Fase III**

### **1.5 Tipo de investigación**

Se diseñó una encuesta estrictamente probabilística, lo que permite estimar con altos niveles de precisión las variables de interés. De acuerdo a la estructura y dinámica de movilidad de los flujos de pasajeros, se diseñaron esquemas de muestreo independientes, tanto para la central de autobuses como para las terminales áreas. El esquema de muestreo utilizado para llevar a cabo la selección de la muestra en cada uno de estos esquemas fue probabilístico, estratificado y bietápico.

## **2. Diseño metodológico**

### **2.1 Marco muestral**

El marco muestral está conformado por los itinerarios de salida de los autobuses de las ciudades Mexicali, Tecate y Tijuana. De igual modo, por el itinerario de salidas de los vuelos nacionales programados en Tijuana y Mexicali en el periodo de referencia de la encuesta.

Debido a que era muy difícil hacer un conteo y aplicación de cédulas filtro piloto que permitiera identificar y ubicar de forma precisa a la población objeto de estudio, se optó por realizar un diseño muestral estratificado en varias etapas, donde la unidad primaria de muestreo es la jornada de trabajo y la unidad última de selección es el turista nacional.

A continuación se enumeran las actividades de campo y de índole administrativa que fue necesario implementar de principio a fin en la solicitud de los itinerarios terrestres y aéreos:





### Observatorio Turístico de Baja California: Fase III

1. Detectar las centrales y líneas de autobuses en las ciudades de Tijuana, Mexicali y Tecate.
2. Detectar de forma precisa las líneas aéreas con destino hacia localidades distintas al lugar de la entrevista.
3. Identificar con claridad las unidades geográficas: zonas y puntos de muestreo.
4. Comunicación y gestión de itinerarios de salida con los ejecutivos de las centrales de autobuses involucradas en el proyecto de investigación.
5. Descarga de los itinerarios de salida de los vuelos nacionales mediante el programa *WorldSpan* y en las páginas web de las aerolíneas.
6. Generar un padrón de unidades de muestreo de las centrales y líneas de autobuses, así como de puertos y líneas aéreas.
7. Clasificación de las unidades geográficas en estratos independientes.
8. Clasificación de las unidades temporales, según día de la semana y mes de referencia.

La primera actividad desarrollada fue realizar un minucioso y exhaustivo trabajo de campo por personal capacitado en determinar de forma apropiada las centrales y líneas de transporte con altas posibilidades de formar parte del marco de muestreo, tanto en las terminales terrestres como en las líneas aéreas que trabajan en los aeropuertos de Tijuana y Mexicali.

Una vez definidas las terminales y líneas de transporte terrestre y aéreo, se identificaron con claridad las unidades geográficas de observación (puertas de salida, salas de espera o de entrada posterior a la revisión), el tipo de sala de espera y el número de puertas en operación (salidas).

Estas actividades permitieron concertar con fineza los espacios determinados para llevar a cabo el levantamiento de cuestionarios, como el procedimiento de selección del entrevistado.



### Observatorio Turístico de Baja California: Fase III

Con las unidades geográficas concertadas y clasificadas, se determinaron las unidades temporales (periodo de referencia de la encuesta), estratificadas por horario, día y mes del levantamiento. Ambos componentes y sus respectivas variables se presentan a continuación:

- **Unidades geográficas.** El espacio geográfico se dividió en tres estratos independientes: localidades, centrales de autobuses y puntos de muestreo.
- **Unidades temporales.** Se dividieron por horario y día de la semana, según el mes de referencia.

La inclusión de la variable espacio y tiempo permite definir con claridad la estructura propia del marco de muestreo y el diseño de muestra utilizado para efectos de la encuesta nacional.

De esta forma se generaron los marcos de muestreo acorde a la dinámica desarrollada en cada punto de levantamiento de encuestas de las ciudades de Baja California: Mexicali, Tecate y Tijuana, según modalidad de entrevista (centrales o aeropuertos). Los resultados de estas actividades, se presentan en los cuadros 1, 2 y 3.

A fin de mostrar el contenido de la estructura de los marcos de muestreo, se considera como ejemplo particular el caso de Tijuana. La información del cuadro 1 en su conjunto denota la estructura y el diseño muestral de la encuesta. Esta estructura consiste de cuatro dominios de estudio perfectamente definidos, donde el primer dominio de estudio alude a la localidad de muestreo como el máximo nivel de agregación de las estimaciones. Enseguida, se tienen a las centrales de autobuses cuya unión da forma a la localidad de muestreo, las cuales funcionan como estratos independientes. Estas centrales contienen a los puntos de muestreo y al interior de estos se localizan los turnos de muestreo. La unión de la localidad, la zona, el punto y el turno de muestreo, representa a la unidad de observación, donde se llevó a cabo



### Observatorio Turístico de Baja California: Fase III

la selección del turista nacional. Para efectos de este estudio éste se intituló como unidad primaria de muestreo.

**Cuadro 1. Marco muestral de la central de autobuses de la ciudad de Tijuana, B.C.**

Localidad	Clave	Zonas	Clave	Puntos	Clave	Turno/Horas	Clave
Tijuana	1	Central de Autobuses	1	Punto Principal	1	00:00 - 08:00	01
Tijuana	1	Central de Autobuses	1	Punto Principal	1	08:00 - 16:00	02
Tijuana	1	Central de Autobuses	1	Punto Principal	1	16:00 - 24:00	03
Tijuana	1	Central de Autobuses	1	Agua caliente	2	06:00 - 14:00	01
Tijuana	1	Central de Autobuses	1	Agua caliente	2	14:00 - 21:00	02
Tijuana	1	Autotransportes Guasave	4	Punto Único	1	08:00 - 16:00	01
Tijuana	1	Autotransportes Guasave	4	Punto Único	1	16:00 - 22:00	02
Tijuana	1	Coordinados Nayarit	5	Punto Único	1	10:00 - 15:00	01
Tijuana	1	Transportes Guamúchil	6	Punto Único	1	11:00 - 19:00	01
Tijuana	1	Tufesa	7	Punto Único	1	07:00 - 15:00	01
Tijuana	1	Tufesa	7	Punto Único	1	16:00 - 23:00	02

**Cuadro 2. Marco muestral de la central de autobuses de la ciudad de Mexicali, B.C.**

Localidad	Clave	Zonas	Clave	Puntos	Clave	Turno/Horas	Clave
Mexicali	2	Central de Autobuses	1	Punto Único	1	08:00 - 16:00	2
Mexicali	2	Central de Autobuses	1	Punto Único	1	16:00 - 24:00	3
Mexicali	2	Autotransportes Guasave	3	Punto Único	1	12:00 - 18:00	1
Mexicali	2	Autotransportes Guasave	3	Punto Único	1	18:00 - 24:00	2
Mexicali	2	Tufesa	4	Punto Único	1	08:00 - 16:00	2
Mexicali	2	Tufesa	4	Punto Único	1	16:00 - 24:00	3
Mexicali	2	Estrellas del Pacífico	5	Punto Único	1	11:00 - 17:00	1
Mexicali	2	Estrellas del Pacífico	5	Punto Único	1	17:00 - 24:00	2
Mexicali	2	Central de Autobuses Turística	6	Punto Único	1	08:00 - 16:00	1
Mexicali	2	Central de Autobuses Turística	6	Punto Único	1	16:00 - 24:00	2



### Observatorio Turístico de Baja California: Fase III

**Cuadro 3. Marco muestral de la central de autobuses de la ciudad de Tecate, B.C.**

Localidad	Clave	Zonas	Clave	Puntos	Clave	Turno/Horas	Clave
Tecate	4	Terminal de Autotransportes Coordinados de Nayarit	2	Punto Único	1	11:00 - 16:00	01
Tecate	4	Terminal de Autotransportes Coordinados de Nayarit	2	Punto Único	1	16:01 - 21:00	02

Concebidos los marcos muestrales, el siguiente paso fue estimar los desplazamientos de personas a nivel de localidad, central de autobús, punto y turno de muestreo con base en la estructura final de los cuadros 1, 2 y 3 con el fin de obtener la información necesaria para el cálculo de las probabilidades de selección. Se considera la ciudad de Tijuana para efectos de mostrar el procedimiento, el resto de las localidades siguió el mismo patrón de cálculo.

#### 2.1.1 Estimación del flujo de pasajeros en centrales de autobuses

La estimación del flujo de personas de las centrales de autobuses se calculó de la siguiente forma:

Donde:

zi:	Las centrales de muestreo (autobuses) de la ciudad de Tijuana, para $i=1, \dots, 5$
tj:	Turnos de una hora, para $j=1, \dots, 24$ , en zi
fk:	La frecuencia de salidas, para $k=1, \dots, n$ , en tj, en zi
$\bar{x}_p$	El promedio de pasajeros en tj, en zi, de acuerdo a fk
dt:	El periodo de levantamiento de la encuesta



### Observatorio Turístico de Baja California: Fase III

El primer cálculo del proceso, consiste en estimar el total (frecuencias) de salidas de autobuses de  $z_i$  en  $t_j$ , independientemente de la línea de autobús, de acuerdo al horario de salida y día de levantamiento (1):

$$\sum_{i=1}^6 z_i \sum_{j=1}^{24} t_j \sum_{k=1}^n f_k \quad (1)$$

Posteriormente, para obtener el flujo de pasajeros por hora efectiva y día de levantamiento, se multiplicó (1) por el promedio de pasajeros; obteniendo de esta manera el flujo total de personas de un día cualquiera de la semana (2):

$$\sum_{i=1}^6 z_i \sum_{j=1}^{24} t_j \sum_{k=1}^n f_k * \bar{x}_p \quad (2)$$

La estimación del total del flujo de pasajeros durante el periodo de levantamiento de cuestionarios se obtuvo multiplicando (2) por el universo de días en muestra, lo cual se muestra en (3):

$$\sum_{i=1}^6 z_i \sum_{j=1}^{24} t_j \sum_{k=1}^n f_k * \bar{x}_p * d_r \quad (3)$$

Para los propósitos de esta encuesta, el periodo de aplicación de cuestionarios fue de acuerdo al número de días efectivos de julio, agosto y septiembre de 2013.

La aplicación del procedimiento anterior permitió estimar el flujo de pasajeros de cada unidad geográfica y temporal según localidad de muestreo, central de autobús, punto y turno de muestreo. Al ser puntos únicos o salidas únicas no fue necesario expandir la información de acuerdo al número de salidas.



## Observatorio Turístico de Baja California: Fase III

### 2.1.2 Estimación del flujo de pasajeros en terminales aéreas

La estimación del flujo de personas en los aeropuertos de Mexicali y Tijuana se realizó bajo el supuesto de que éstos son únicos, es decir, cada ciudad sólo cuenta con una única terminal aérea. De igual modo, para mostrar el procedimiento se considera a la ciudad de Tijuana, ya que para Mexicali se siguió la misma estrategia; el cálculo se muestra a continuación:

Donde:

t <sub>j</sub> :	Jornadas de trabajo de 8 horas, para j=1, 2, 3
f <sub>k</sub> :	La frecuencia de vuelos, para k=1,.. n, en t <sub>j</sub> ,
C <sub>p</sub> :	Capacidad potencial de pasajeros de los vuelos nacionales
dt:	Periodos de un mes para la ciudad de referencia

El primer cálculo del proceso consiste en estimar el total (frecuencias) de salidas de vuelos en t<sub>j</sub>, independientemente de la línea aérea, de acuerdo al horario de salida y día de levantamiento (1):

$$\sum_{j=1}^{24} t_j \sum_{k=1}^n f_k \quad (1)$$

Posteriormente, para obtener el total de vuelos, como el total de desplazamientos por hora efectiva y día de levantamiento, se multiplicó (1) por el periodo de referencia, obteniendo de esta manera el total de vuelos de un día cualquiera de la semana (2):

$$\sum_{j=1}^{24} t_j \sum_{k=1}^n f_k * C_p * d_t \quad (2)$$

La aplicación del procedimiento anterior permitió estimar el total del flujo de personas de cada unidad geográfica y temporal, según localidad de muestreo, terminal aérea,



## Observatorio Turístico de Baja California: Fase III

punto y turno de muestreo. Al ser puntos únicos o salidas únicas no fue necesario expandir la información de acuerdo al número de salidas.

### 2.2 Estratificación del espacio geográfico y cálculo de probabilidades de primer orden

#### 2.2.1 Centrales de autobuses

En la siguiente etapa se desarrolla el procedimiento utilizado para determinar las probabilidades de inclusión de primer orden con los criterios establecidos en el diseño muestral del proyecto de investigación.

Donde:

$N_h$	Flujo de pasajeros en la h-ésima zona de muestreo para $h=1, \dots, L$ .
$N = \sum_{h=1}^L N_h$	Total del flujo de pasajeros en la h-ésima zona de muestreo
$M_{jh}$	Flujo de pasajeros del j-ésimo punto de muestreo, de la h-ésima zona de muestreo.
$B_{ijh}$	Flujo de pasajeros del i-ésimo turno de muestreo, del j-ésimo punto de muestreo, de la h-ésima zona de muestreo.

#### Primera partición del espacio geográfico: localidad de muestreo

La primera partición del espacio geográfico inicia con el máximo nivel de agregación o inferencia, que para efectos de este estudio está representada por la localidad de muestreo, por lo que la probabilidad de inclusión de primer orden es la siguiente:

$$\pi_i = \frac{N}{N} = 1$$

Como la localidad se incluye con certeza, la probabilidad de inclusión de primer orden es 1.



## Observatorio Turístico de Baja California: Fase III

### Segunda partición del espacio geográfico: zonas de muestreo

Para determinar la probabilidad de inclusión de primer orden de la zona de muestreo se procedió de la siguiente forma:

$$\pi_h = \frac{N_h}{N}$$

Donde  $\pi_h$  es la probabilidad de inclusión de primer orden de la h-ésima zona de muestreo.

### Tercera etapa de partición del espacio geográfico: puntos de muestreo

La probabilidad de inclusión de primer orden del punto de muestreo viene dada de la siguiente forma:

$$\pi_{jh} = \frac{M_{jh}}{N_h}$$

Donde  $\pi_{jh}$  es la probabilidad de inclusión de primer orden del j-ésimo punto de muestreo, de la h-ésima zona de muestreo.

### Cuarta etapa de partición: turnos de muestreo

La probabilidad inclusión de primer orden de esta etapa viene dada por:

$$\pi_{ijh} = \frac{B_{ijh}}{M_{jh}}$$

Donde  $\pi_{ijh}$  es la probabilidad de inclusión de primer orden del i-ésimo turno de muestreo, del j-ésimo punto de muestreo, de la h-ésima zona de muestreo.





### Observatorio Turístico de Baja California: Fase III

#### Probabilidad inclusión de primer orden de la UPM

La probabilidad total de inclusión de primer orden de cada unidad de observación resulta multiplicando factor a factor las probabilidades de inclusión de cada etapa, la cual viene dada de la siguiente forma:

$$\pi_{lhji} = \pi_l \times \pi_h \times \pi_{jh} \times \pi_{ijh}$$

Donde,  $\pi_{lhji}$  representa la probabilidad de inclusión de la unidad primaria de muestreo, enfatizando que para efectos de la presente investigación la probabilidad total de inclusión de primer orden es proporcional al número de elementos presentes en cada subconjunto de estudio.

#### 2.2.2 Terminales aéreas

En la siguiente etapa se muestra el procedimiento utilizado para determinar las probabilidades de inclusión de primer orden en las terminales aéreas.

Donde:

$N_h$	Flujo de personas en la h-ésima zona de muestreo para $h=1, \dots, L$ .
$N = \sum_{h=1}^L N_h$	Total del flujo de personas en la h-ésima zona de muestreo.
$M_{jh}$	Flujo de personas en el j-ésimo punto de muestreo, de la h-ésima zona de muestreo.
$B_{ijh}$	Flujo de personas del i-ésimo turno de muestreo, del j-ésimo punto de muestreo, de la h-ésima zona de muestreo.



## Observatorio Turístico de Baja California: Fase III

### Primera partición del espacio geográfico: localidad de muestreo

La primera partición del espacio geográfico inicia con el máximo nivel de agregación o inferencia, que para efectos de este estudio está representada por la localidad de muestreo, por lo que la probabilidad de inclusión de primer orden es la siguiente:

$$\pi_l = \frac{N}{N} = 1$$

Como la localidad se incluye con certeza, la probabilidad de inclusión de primer orden es 1.

### Segunda partición del espacio geográfico: zonas de muestreo

Para determinar la probabilidad de inclusión de primer orden de la zona de muestreo se procedió de la siguiente forma:

$$\pi_h = \frac{N_h}{N}$$

Al ser única la zona de muestreo, se tiene que  $N_h = N$ , por lo que la probabilidad de inclusión de primer orden de la h-ésima zona de muestreo  $\pi_h$  es igual a 1.

### Tercera etapa de partición del espacio geográfico: puntos de muestreo

La probabilidad de inclusión de primer orden del punto de muestreo viene dada de la siguiente forma:

$$\pi_{jh} = \frac{M_{jh}}{N_h}$$

Donde  $\pi_{jh}$  es la probabilidad de inclusión de primer orden del j-ésimo punto de muestreo, de la h-ésima zona de muestreo.



## Observatorio Turístico de Baja California: Fase III

### Cuarta etapa de partición: turnos de muestreo

La probabilidad inclusión de primer orden de esta etapa viene dada por:

$$\pi_{ijh} = \frac{B_{ijh}}{M_{jh}}$$

Donde  $\pi_{ijh}$  es la probabilidad de inclusión de primer orden del i-ésimo turno de muestreo, del j-ésimo punto de muestreo, de la h-ésima zona de muestreo.

### Probabilidad inclusión de primer orden de la UPM

Por lo tanto, la probabilidad de inclusión de primer orden de cada unidad de observación resulta multiplicando factor a factor las probabilidades de inclusión de cada etapa, la cual viene dada de la siguiente forma:

$$\pi_{lhji} = \pi_l \times \pi_h \times \pi_{jh} \times \pi_{ijh}$$

Donde  $\pi_{lhji}$  representa el peso de la unidad primaria de muestreo, enfatizando que para efectos de la presente investigación la probabilidad total de inclusión de primer orden es proporcional al número de elementos presentes en cada subconjunto de estudio.

### 2.3 Tamaño de muestra

La información empleada para determinar el tamaño de muestra se restringe al tamaño del flujo de personas en centrales de autobuses y terminales aéreas que se estimó en los periodos de referencia, **como a la precisión requerida para los parámetros prioritarios de la encuesta, el presupuesto disponible y los niveles de inferencia a los que se requiere que sean representativos los resultados.**



### Observatorio Turístico de Baja California: Fase III

La expresión empleada para determinar el tamaño de muestra se presenta a continuación:

$$n \geq \frac{(Z_{\alpha/2})^2}{\epsilon^2} * deff * tnr$$

Donde:

n	Tamaño de muestra
p	Máxima varianza del parámetro desconocido
e	Nivel aceptable de error, en porcentaje
z	Variable normal estándar para el nivel de confianza seleccionado
deff	Efecto de diseño
tnr	Tasa de no respuesta

Se considera un error de muestreo máximo permitido del 5.0 % (e=0.05) con un valor de Z=1.96 que corresponde a un nivel de confianza del 95 %. Se modifica la muestra por el efecto de diseño (2.25) y por la tasa de no respuesta, 15 %. Se parte de un tamaño de muestra requerido para un muestreo aleatorio simple, pero al tratarse de un muestreo estratificado en varias etapas se modifica para reflejar la utilización de un diseño distinto al aleatorio simple por el efecto de diseño. El efecto de diseño resulta del cociente de la varianza del parámetro estimado del esquema de muestreo distinto al aleatorio de tamaño **n**, entre la varianza del mismo parámetro resultante de un muestreo aleatorio simple, del mismo tamaño **n**, o sea:

$$deff = \frac{V_{diferente}}{V_{mas}}$$

La información utilizada para el cálculo del deff y de la tnr, proviene de la encuesta de “Caracterización de los visitantes nacional en las ciudades fronterizas de Baja California: Mexicali y Tijuana (invierno 2012-2013)”, aplicada en los meses de noviembre y diciembre de 2012 y enero de 2013.



## Observatorio Turístico de Baja California: Fase III

### 2.3.1 Afijación de la muestra

La afijación de la muestra se realizó de manera proporcional a los tamaños de los estratos  $N_h$ , mediante la siguiente expresión:

$$n_h = n \frac{N_h}{N} = nW_h$$

### 2.3.2 Selección de la muestra

El proceso de selección de la muestra incluye en su conjunto a las unidades geográficas y temporales. Dicha selección, parte de la matriz de probabilidades de selección de primer orden.

De esta forma  $\pi_{lhji}$  es la probabilidad de que el elemento  $v_{lhji}$  esté representado en la muestra, y se denomina probabilidad de inclusión de primer orden. Como  $v_{lhji}$  puede pertenecer o no a la muestra, esto queda representado como:

$$m_{lhji} : S \rightarrow \{0,1\}$$

De tal forma:

$$m_{lhji}(s) = 1, \text{ si } v_{lhji} \in s \text{ y } m_{lhji}(s) = 0, \text{ si } v_{lhji} \notin s \quad \forall s \in S, \forall v_{lhji} \in \Omega$$

Por lo tanto,  $m_{lhji}$  es una variable aleatoria definida sobre S, la cual tiene una distribución de probabilidad igual a:

$$P[m_{lhji} = 1] = \sum_{s \in S; v_{lhji} \in s} P(s) = \pi_{lhji}$$
$$P[m_{lhji} = 0] = 1 - \pi_{lhji}$$



## Observatorio Turístico de Baja California: Fase III

La selección de la muestra se realizó mediante el programa estadístico SPSS (IBM SPSS Statistics 20), partiendo de la matriz de probabilidades de selección denominada como  $\pi_{lhji}$ , la cual concentra los pesos acumulados del marco muestral y que representan a toda la cobertura del universo geográfico y temporal.

### 2.4 Mediciones

Una vez que las jornadas de trabajo han sido seleccionadas, éstas se proporcionaron a la coordinación de trabajo de campo para que se distribuyera entre los supervisores de las localidades presentes en el proyecto de investigación.

En el espacio geográfico seleccionado para llevar a cabo el levantamiento de cuestionarios se definieron los siguientes criterios:

- **Flujos.** Una vez *in situ* en el punto y jornada a medir, el personal de campo tomó una muestra sistemática del flujo observado y a la vez llevaba un registro del flujo total entre cada persona del flujo seleccionado de manera sistemática. Para ello en cada punto-jornada trabajaron al menos dos personas, una que llevaba el registro del flujo total y selección sistemática y otra que realizaba las entrevistas. Durante un mismo turno, por regla, cada pareja debía alternarse cada hora para realizar una u otra actividad. El registro del flujo se mantuvo continuo durante la jornada de trabajo.
- **Filtro.** Distinguía entre los interceptados por la muestra sistemática, los desplazamientos de interés de los que no lo eran. Era aplicado por quien estaba asignado para realizar entrevistas. El filtro constaba de unas preguntas para diferenciar la población de interés dentro del flujo total.
- **Entrevista.** Una vez que algún interceptado pasara el filtro, se le aplicó la entrevista dirigida al flujo de interés en el punto-jornada, y así sucesivamente con otros interceptados.



## Observatorio Turístico de Baja California: Fase III

### 2.4.1 Proceso de estimación

El factor de expansión se calculó de acuerdo a la unidad primaria de muestreo seleccionada, que se denomina como  $\pi_{l_{hji}}$ , día de la semana y cuestionario. Este factor de expansión resulta de tres multiplicadores: del factor de discriminación, del factor geográfico y del factor temporal. El proceso de estimación siguió exactamente las pautas establecidas en el diseño muestral y selección de la muestra.

**Cabe destacar que utilizó el mismo procedimiento de estimación tanto en las centrales de autobuses como en las terminales aéreas.**

### 2.4.2 Factor de discriminación

El primer elemento se refiere al factor de discriminación, el cual se obtuvo directamente de las labores de campo y correspondía al desplazamiento total de personas ( $n$ ) que pasan por un determinado punto de muestreo seleccionado para la muestra, aplicando la cédula filtro y el cuestionario, y  $r$  es el número de renglones que se ocuparon de la cédula filtro hasta encontrar a un visitante nacional, por lo que el factor de la etapa viene dado de la siguiente forma:

$$\left(\frac{r}{n}\right)^{-1}$$

De esta manera, se obtuvo la probabilidad de seleccionar un turista nacional en centrales de autobuses y en terminales aéreas.

En el caso de cuestionarios incompletos o con presencia de información falseada, se imputó  $n$  y la probabilidad de inclusión de primer orden  $\pi_{l_{hji}}$  de acuerdo a la fecha,



## Observatorio Turístico de Baja California: Fase III

localidad, zona, punto y turno, al resto de cuestionarios para seguir manteniendo la representación obtenida en campo. Posteriormente se eliminaron estos cuestionarios.

### 2.4.3 Factor geográfico

Una vez determinado el primer multiplicador, se procedió con el cálculo del factor de las unidades geográficas:

#### Localidad de muestreo.

Donde  $\pi_L = \{\text{las localidades de muestreo}\} = \{\pi_1, \pi_2, \dots, \pi_c\}$ , que corresponde a la probabilidad de primer orden de las ciudades de muestreo, respecto al flujo total de personas que se desplazan por dicha ciudad. Al ser la primera etapa de selección, se maneja como dominio de estudio, por lo que su probabilidad de selección es 1 y se define como:

$$(\pi_L)^{-1}$$

#### Zonas de muestreo.

Donde  $\pi_H = \{\text{las zonas de muestreo de la ciudad F}\} = \{\pi_1, \pi_2, \dots, \pi_c\}$ , que representan la probabilidad de inclusión de primer orden de la zona en la ciudad de muestreo respecto al flujo que se desplaza por la zona de muestreo. Estas probabilidades se calculan con la información estimada del flujo de personas. Si además se tiene que  $\pi_j$  es la zona seleccionada en la ciudad p  $\pi_H$ , la probabilidad asociada con las zonas de muestreo de la ciudad H es:

$$(p\pi_H)^{-1}$$





## Observatorio Turístico de Baja California: Fase III

### Punto de muestreo.

Donde  $\pi_j = \{\text{los puntos de muestreo de la zona G}\} = \{\pi_1, \pi_2, \dots, \pi_c\}$ , que representan la probabilidad de inclusión de primer orden de cada punto de muestreo de su respectiva zona en función del flujo que se desplaza por dicho punto. Las probabilidades se calcularon con la estimación del flujo. Si además  $S_i$  es el punto de muestreo seleccionado y se denota como  $p\pi_j$  la probabilidad de selección de dicho punto, el factor asociado con los puntos de muestreo se determina como:

$$(p\pi_j)^{-1}$$

### Turnos de muestreo.

Sea  $\pi_l = \{\text{los turnos de muestreo del día K}\} = \{\pi_1, \pi_2, \dots, \pi_c\}$ , que representa la probabilidad de inclusión de primer orden de cada turno en el punto o zona de muestreo. Sea  $p\pi_i$  la probabilidad de selección del turno en el día. De esta forma el factor asociado al turno de muestreo es:

$$(p\pi_l)^{-1}$$

Por lo que la probabilidad de inclusión de primer orden de la unidad de observación geográfica (UPM) seleccionada es:

$$(\pi_L \times p\pi_H \times p\pi_j \times p\pi_l)^{-1}$$

## 2.5 Factor temporal

**Días de muestreo.** Se muestra el procedimiento en particular para un día de la semana.

Donde  $D_M = \{\text{lunes del periodo de referencia de la encuesta}\} = \{m_1, m_2, \dots, m_k\}$ , de tal forma que  $nD_M = \{\text{lunes seleccionados del periodo de referencia de la encuesta}\} =$



### Observatorio Turístico de Baja California: Fase III

$\{m_1, m_2, \dots, m_k\}$ . De esta manera el factor de expansión de esta variable queda determinado como:

$$P_i = \left( \frac{nD_M}{D_M} \right)^{-1}$$

A la probabilidad de inclusión determinada en el procedimiento anterior, se le realiza un ajuste por días faltantes. Un día faltante se define como aquel día que por alguna razón justificable no se levantó ningún cuestionario durante la jornada de trabajo.

Por lo tanto, al restar al número de días que cayeron en muestra el número de días faltantes se obtiene el número de días en las que fueron seleccionadas y levantadas las jornadas de trabajo.

De este forma, sea **dm (D(Ham))** el número de días faltantes de los lunes seleccionados del mes de referencia.

Lo anterior se reduce a lo siguiente:

**Factor de ajuste:**

$$d_{af}(D(H_{am})) = nD_M - d_m(D(H_{am}))$$

Por lo que la probabilidad del factor temporal queda definida de la siguiente manera:

**Factor temporal:**

$$P'_i = \frac{D_M}{d_{af}(D(H_{am}))}$$

De esta forma queda determinada la probabilidad de inclusión del factor temporal.



### Observatorio Turístico de Baja California: Fase III

Por lo que el factor de expansión para un determinado punto de muestreo y cuestionario queda determinado como:

$$\left( \left[ \left( \frac{k}{n} \right)^{-1} \right] * (\pi_L \ x \ p \pi_H \ x \ p \pi_J \ x \ p \pi_I)^{-1} * P'_i \right)$$



## Observatorio Turístico de Baja California: Fase III

### Referencia:

Bringas Rábago, Nora L. (Coord.) (2013). "Caracterización de los visitantes nacionales en las ciudades fronterizas de Baja California: Mexicali y Tijuana (invierno 2012-2013)", Tijuana. Reporte de investigación. El Colegio de la Frontera Norte (Cofef)-Secretaría de Turismo del Estado de Baja California (Secture).