

Los casos de SIDA diagnosticados en España en los últimos años vienen recogidos en la siguiente tabla, clasificados por grupo de riesgo del paciente

	1993	1994	1995	1996	1997
<i>Por drogas inyectables</i>	3373	4723	4432	4203	3143
<i>Heterosexuales</i>	631	952	1044	1113	950
<i>Homosexuales/bisexuales</i>	778	898	934	766	580
<i>Drogas inyect. y homosexuales</i>	68	92	81	73	61
<i>Hijos de madre de riesgo</i>	82	76	79	46	37
<i>Por hemoderivados</i>	56	66	52	44	25
<i>Por transfusiones</i>	32	23	21	14	12
<i>Otros</i>	321	291	302	249	290
	5341	7121	6945	6508	5098

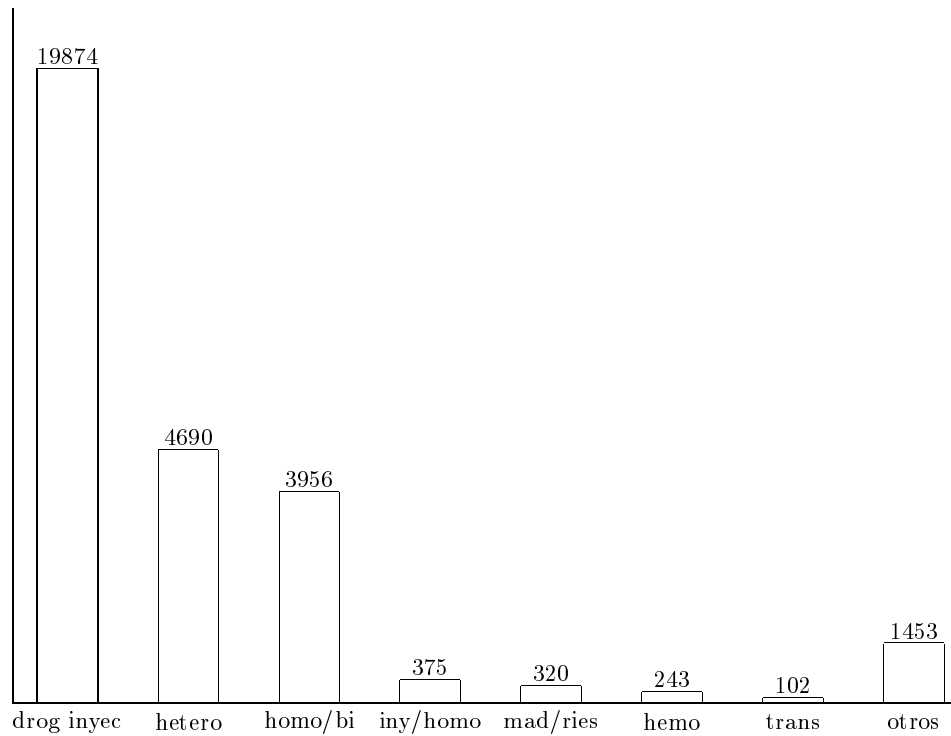
Analizar descriptivamente estos datos.

Se trata de unos datos correspondientes a una distribución bidimensional de frecuencias. Para este tipo de datos (CB-sección 2.4) lo primero que deberemos hacer es identificar de qué tipo son los caracteres que dan origen a la *tabla de doble entrada* donde aparece la distribución bidimensional de frecuencias.

En este ejercicio, uno de los caracteres, *grupo de riesgo*, es de tipo cualitativo. La *distribución (de frecuencias absolutas) marginal* de dicho carácter, se obtiene directamente de la tabla de contingencia, calculando, para cada modalidad del carácter, el número de individuos que presentan dicha modalidad a lo largo de todas las modalidades del otro carácter; es decir,

	n_i
<i>Por drogas inyectables</i>	19874
<i>Heterosexuales</i>	4690
<i>Homosexuales/bisexuales</i>	3956
<i>Drogas inyect. y homosexuales</i>	375
<i>Hijos de madre de riesgo</i>	320
<i>Por hemoderivados</i>	243
<i>Por transfusiones</i>	102
<i>Otros</i>	1453
	31013

El histograma de dicha distribución de frecuencias absolutas es



De las medidas de posición o dispersión, sólo podemos decir que el grupo de riesgo más frecuente, es decir la *moda*, es el, *ser consumidor de drogas inyectables*.

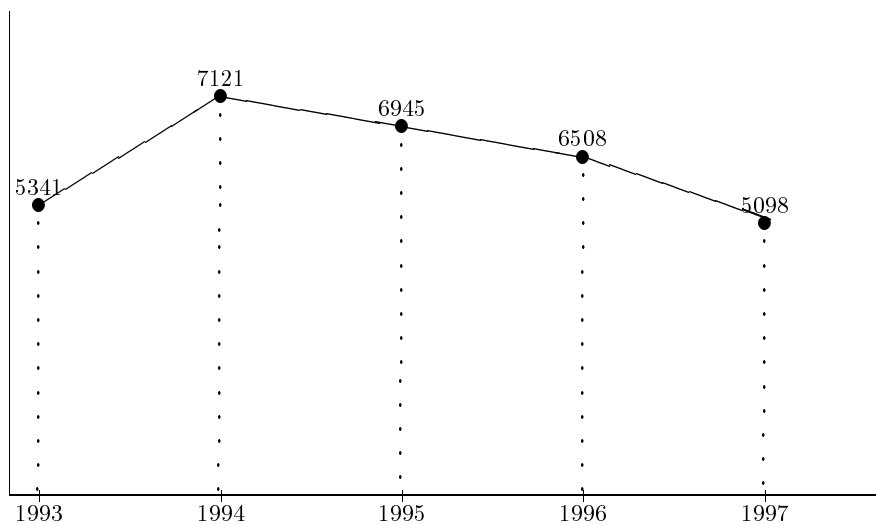
Respecto al otro carácter, *año*, aunque aparentemente es de tipo cuantitativo, no van a tener tampoco mucho sentido las medidas de posición o dispersión que resuman la distribución marginal correspondiente, la cual sería

	n_i
1993	5341
1994	7121
1995	6945
1996	6508
1997	5098
	31013

Las modalidades del carácter *año* actúan más como *etiquetas* que como valores numéricos. No obstante, presentan una peculiaridad muy interesante: van ordenadas. Esto las diferencia de los caracteres cualitativos.

Además, el que ese carácter se refiera al tiempo, hace que el estudio de este tipo de datos sea especialmente importante y diferenciado de los que hasta ahora hemos considerado.

A la vista de la distribución de frecuencias marginal anterior, ha habido una evolución temporal primero creciente, con un máximo en 1994, para luego ir disminuyendo el número de casos, lo que parece indicar una *tendencia* decreciente en el número de individuos con la enfermedad, como muestra el gráfico siguiente



Pero, posiblemente, tienen más interés las *distribuciones condicionadas*, por años, en especial en términos relativos

Grupo de riesgo/1993	n_i	f_i
Por drogas inyectables	3373	0'6315
Heterosexuales	631	0'1181
Homosexuales/bisexuales	778	0'1457
Drogas inyect. y homosexuales	68	0'0127
Hijos de madre de riesgo	82	0'0153
Por hemoderivados	56	0'0105
Por transfusiones	32	0'0060
Otros	321	0'0602
	5341	1

Grupo de riesgo/1994	n_i	f_i
<i>Por drogas inyectables</i>	4723	0'6632
<i>Heterosexuales</i>	952	0'1337
<i>Homosexuales/bisexuales</i>	898	0'1261
<i>Drogas inyect. y homosexuales</i>	92	0'0129
<i>Hijos de madre de riesgo</i>	76	0'0107
<i>Por hemoderivados</i>	66	0'0093
<i>Por transfusiones</i>	23	0'0032
<i>Otros</i>	291	0'0409
	7121	1

Grupo de riesgo/1995	n_i	f_i
<i>Por drogas inyectables</i>	4432	0'6382
<i>Heterosexuales</i>	1044	0'1503
<i>Homosexuales/bisexuales</i>	934	0'1345
<i>Drogas inyect. y homosexuales</i>	81	0'0117
<i>Hijos de madre de riesgo</i>	79	0'0114
<i>Por hemoderivados</i>	52	0'0075
<i>Por transfusiones</i>	21	0'0030
<i>Otros</i>	302	0'0434
	6945	1

Grupo de riesgo/1996	n_i	f_i
<i>Por drogas inyectables</i>	4203	0'6458
<i>Heterosexuales</i>	1113	0'1710
<i>Homosexuales/bisexuales</i>	766	0'1177
<i>Drogas inyect. y homosexuales</i>	73	0'0112
<i>Hijos de madre de riesgo</i>	46	0'0071
<i>Por hemoderivados</i>	44	0'0068
<i>Por transfusiones</i>	14	0'0022
<i>Otros</i>	249	0'0382
	6508	1

Grupo de riesgo/1997	n_i	f_i
<i>Por drogas inyectables</i>	3143	0'6165
<i>Heterosexuales</i>	950	0'1863
<i>Homosexuales/bisexuales</i>	580	0'1138
<i>Drogas inyect. y homosexuales</i>	61	0'0120
<i>Hijos de madre de riesgo</i>	37	0'0073
<i>Por hemoderivados</i>	25	0'0049
<i>Por transfusiones</i>	12	0'0024
<i>Otros</i>	290	0'0568
	5098	1

De hecho se podría decir que la tabla de doble entrada del enunciado tiene una interpretación diferente de la habitual considerada en CB-sección 2.4. Aquí, no es tanto que cada uno de los $n = 31013$ enfermos fuera asignado a un grupo de riesgo y a un año, hasta formar la tabla de contingencia del enunciado, sino que, muy posiblemente, se fueran recogiendo datos, año a año (es decir, se fueran construyendo las distribuciones condicionadas anteriores), para formar finalmente la tabla de datos del enunciado.

Tantas son las particularidades de este tipo de datos, en los que una de las variables es el tiempo, que existe un Método Estadístico Aplicado específico para su tratamiento, el cual recibe el nombre de *Análisis de Series Temporales*.

Analizando la evolución temporal condicionada por el grupo de riesgo, es decir, las ocho distribuciones de frecuencias absolutas de la variable *tiempo* condicionada por cada modalidad de la variable *grupo de riesgo*, obtendríamos las siguientes distribuciones de frecuencias (sólo damos las dos primeras)

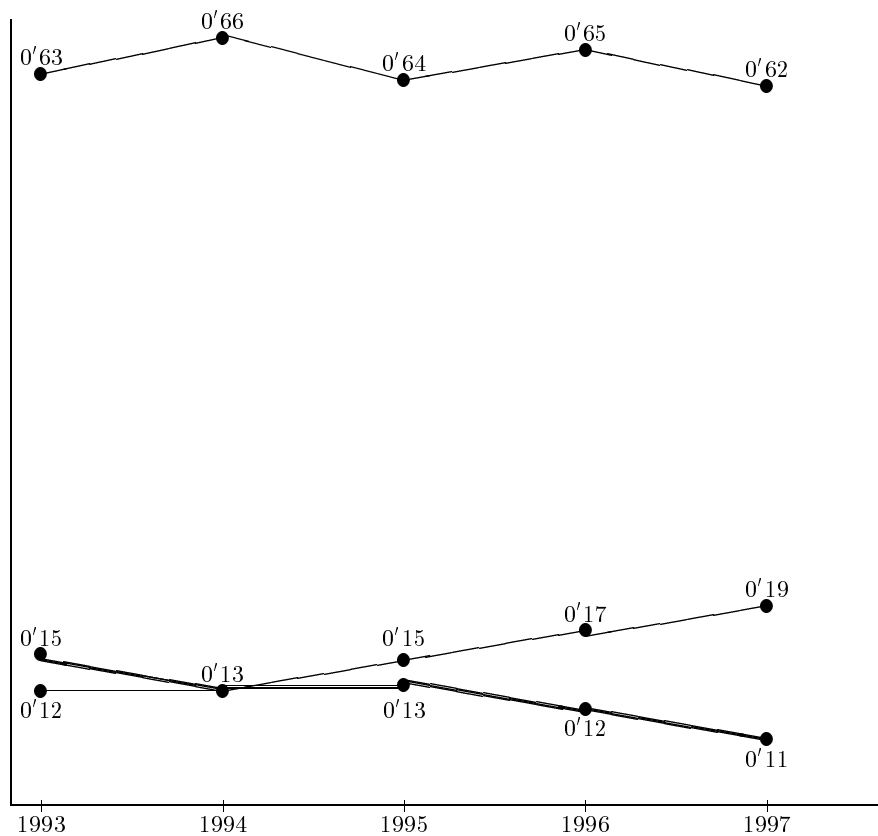
Años/Drogas inyectables	n_i	Años/Heterosexuales	n_i
<i>1993</i>	3373	<i>1993</i>	631
<i>1994</i>	4723	<i>1994</i>	952
<i>1995</i>	4432	<i>1995</i>	1044
<i>1996</i>	4203	<i>1996</i>	1113
<i>1997</i>	3143	<i>1997</i>	950
	19874		4690

Pero posiblemente, lo más interesante de todo sea analizar de forma global, y mediante una representación gráfica, la evolución temporal de los grupos de riesgo en porcentaje; es decir, las distribuciones relativas condicionadas por años, antes calculadas y que permiten realizar los siguientes gráficos, en donde hemos omitido la evolución de la modalidad *otros*.

Hemos realizado dos gráficos para clarificar los dibujos y, sobre todo, para poder variar la escala de las ordenadas y apreciar así con más nitidez las diferencias en la evolución de las frecuencias.

Así, aunque la evolución temporal de la distribución marginal de los años, antes representada, indicaba una evolución con un máximo en 1994 para luego mantenerse una tendencia descendente, del gráfico que sigue se deduce que el porcentaje, sobre el total de casos registrados cada año, de enfermos cuya causa de contagio fue la de *por drogas inyectables*, se mantiene

prácticamente constante; que el grupo de *heterosexuales* crece, y que el de *homosexuales/bisexuales* decrece.



Por último, en el gráfico siguiente, se observa que el grupo de riesgo de personas que consumen *drogas inyectables* y son *homosexuales*, manifiesta una tendencia casi constante, que el grupo de *hijos con madre de riesgo* disminuye a rachas, es decir, con períodos intermedios constantes; que el grupo de contagio *por hemoderivados* decrece muy significativamente ya desde niveles muy bajos, y que, por último, el grupo de *transfusiones* prácticamente (o mejor dicho, porcentualmente respecto al número total de casos) casi ha desaparecido.

