Université Paris 1 Panthéon Sorbonne,

Institut de démographie

Cours d'analyse démographique niveau : **Master de démographie** par Alexandre Avdeev,

Chapitre 4

Analyse des taux : comparaison, standardisation, décomposition

1. Equation du bilan démographique avec les taux bruts
2. Relation entre les taux bruts et les taux par âge, difficultés de la comparaison des taux bruts.
3. Méthodes de calculs des taux « comparatifs », ou la standardisation
   - Directe
   - Indirecte
   - Autre possibilité de la comparaison : la standardisation inverse
4. Décomposition d'une différence entre les taux bruts
   - Modèles additifs sans et avec interaction.
   - Modèles multiplicatifs

rappel

Équation de bilan démographique avec des valeurs absolues

$$P_t - P_0 = N_{0,t} - D_{0,t} + I_{0,t} - E_{0,t}$$

P_t – nombre de survivants au moment t

P_0 – nombre de survivants au moment 0 (précédant à t)

$N_{0,t}$ – nombre de naissances durant la période entre 0 et t

$D_{0,t}$ – nombre de décès durant la période entre 0 et t

$I_{0,t}$ – nombre migrants arrivés durant la période entre 0 et t (effet d'immigration)

$E_{0,t}$ – nombre migrants partis durant la période entre 0 et t (effet d'émigration)

	Population au 1 janvier 2000	au cours de l'année			Population au 1 janvier 2001	Ajustement statistique
		Naissance	Décès	solde migratoire		
France	58 858 198	(+)774 782	(-)535 066	(+) 70 000	(=) 59 266 572	(+) 94 456
Danemark	5 330 020	(+) 67 084	(-) 57 986	(+) 10 094	(=) 5 349 212	(+) 596

C'est un bon instrument pour les estimations nationales, mais les comparaisons internationales sont difficiles avec des valeurs absolues rappel

Équation de bilan démographique avec les taux bruts

$$\frac{P_t - P_0}{t \cdot \overline{P}_{0,t}} = \frac{N_{0,t}}{t \cdot \overline{P}_{0,t}} - \frac{D_{0,t}}{t \cdot \overline{P}_{0,t}} + \frac{I_{0,t}}{t \cdot \overline{P}_{0,t}} - \frac{E_{0,t}}{t \cdot \overline{P}_{0,t}}$$

$$TBA = (TBN_{0,t} - TBD_{0,t}) + (TBI_{0,t} - TBE_{0,t})$$

$$TBA = TBAN_{0,t} + TBAM_{0,t}$$

Années vécues par la population durant la période T(0,t), ou « population exposée » (personnes-années)

France ➔ 0.5 x (58 858 198 + 59 266 572) = 59 062 385

Danemark ➔ 0.5 x (5 330 020 + 5 349 212) = 5 339 616

	TBA	TBN	TBM	TBAM
France	0, 006914	(+) 0,0132	(-) 0,00899	(+) 0,00119
Danemark	0,00359	(+) 0,01256	(-) 0,01086	(+) 0,00189

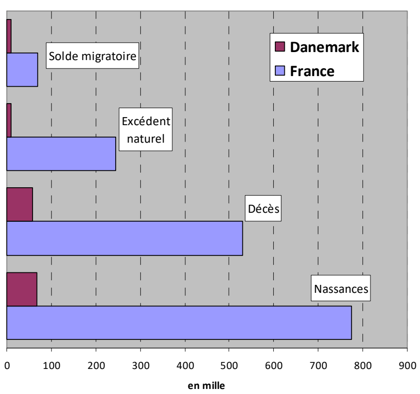
ou (pour facilité la perception) pour 1000 population (pour 1000 années vécues)

	TBA	TBN	TBM	TBAM
France	6,914	(+) 13,118	(-) 8,988	(+) 1,185
Danemark	3,594	(+) 12,563	(-) 10,860	(+) 1,779

rappel

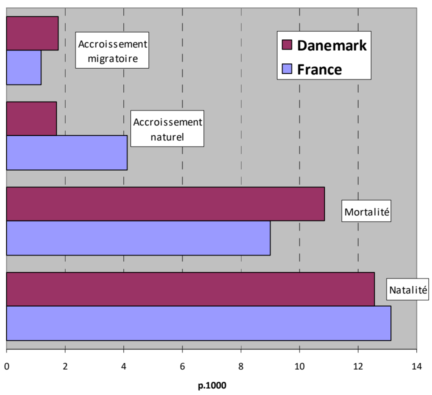
Analyse visuelle des composants du mouvement de la population en France et en Danemark à l'an 2000

A. Nombre d'événements



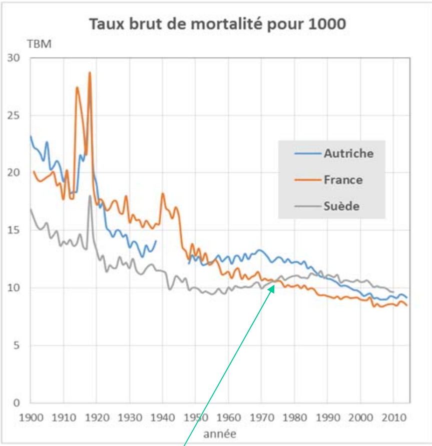
C'est difficile à interpréter à cause de la domination numérique de la population française

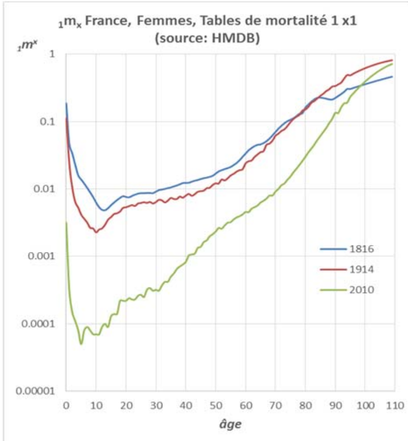
B. Taux bruts



C'est facile à interpréter puisque l'effectif de la population est réduit à 1000 pour les deux pays

Comment comparer les niveaux et interpréter la dynamique des taux bruts ?





Pourquoi le TBM en Suède > TBM en France ?

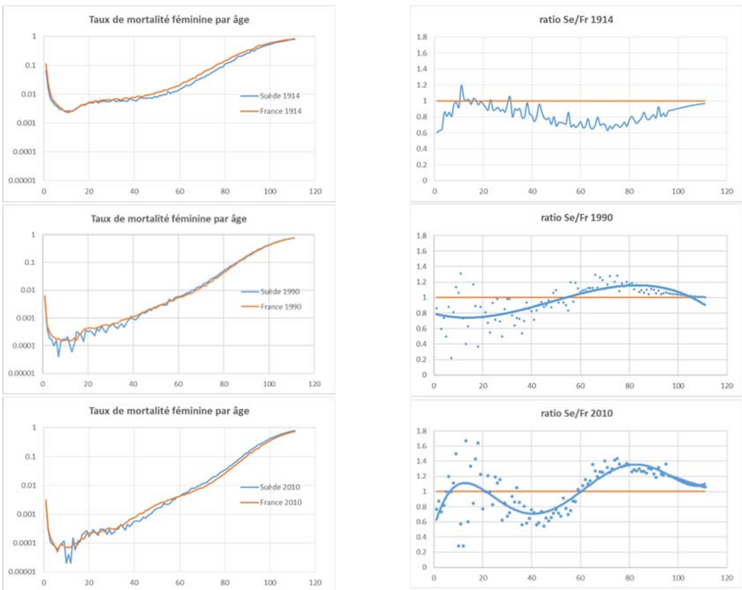
En 1980 e_o,	femmes	homme :
Autriche	76.0	69.0
France	78.9	70.2
Suède	78.0 (-0.9)	72.8 (+2.6)

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

5

5

Comparaison visuelle de la mortalité féminine par âge en France et en Suède en 1914, 1990 et 2010 (n'est pas vraiment satisfaisante)



Source: Human Mortality Data Base (consultation 6/11/2016)

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

6

6

2. Relation entre les taux bruts et les taux par âge, difficultés de la comparaison des taux

Rapport entre les taux par âge et les taux bruts

Soit $_nD_x$ le nombre de décès à l'intervalle d'âge [***x, x+n***) durant une période entre ***0*** et ***t***;
$_nB_x$ le nombre de naissance à l'âge [***x, x+n***) durant une période entre ***0*** et ***t***.

taux par âge

événement par âge

$$_n m_x = \frac{_nD_x}{t \cdot {_n\overline{P}_x}}$$ → **taux de mortalité pour âge « x » et période [0,t)** → $$_nD_x = {_nm_x} \cdot t \cdot {_n\overline{P}_x}$$

$$_n f_x = \frac{_nB_x}{t \cdot {_n\overline{P}_x}}$$ → **taux de fécondité pour âge « x » et période [0,t)** → $$_nB_x = {_nf_x} \cdot t \cdot {_n\overline{P}_x}$$

Alors les sommes d'événements et, par conséquent, les taux bruts, dépendent des taux par âge

Etant donné que la durée de la période est une année (t = 1), nous pouvons nous passer du symbole t des formules

Les taux bruts sont des moyennes arithmétiques des taux par âge (démonstration)

$$TBM = \frac{D}{t \cdot \bar{P}} = \frac{\sum_x {}_nD_x}{t \cdot \sum_x {}_n\bar{P}_x} = \frac{t \cdot \sum_x {}_nm_x \cdot {}_n\bar{P}_x}{t \cdot \sum_x {}_n\bar{P}_x} = \sum_x {}_nm_x \cdot \frac{{}_n\bar{P}_x}{\sum_x {}_n\bar{P}_x} = \sum_x {}_nm_x \cdot {}_np_x$$

avec ${}_np_x$ - élément de structure de la population par âge (**proportion**)

$$TBN = \frac{B}{t \cdot \bar{P}} = \frac{\sum_x {}_nB_x}{t \cdot \sum_x {}_n\bar{P}_x} = \frac{t \cdot \sum_x {}_nf_x \cdot {}_n\bar{P}_x}{t \cdot \sum_x {}_n\bar{P}_x} = \sum_x {}_nf_x \cdot \frac{{}_n\bar{P}_x}{\sum_x {}_n\bar{P}_x} = \sum_x {}_nf_x \cdot {}_np_x$$

sous une forme générale: ⇒

$$TB = \sum_{x=0}^{\omega} \tau_x \cdot w_x$$

avec ***x*** – âge
τ_x – taux par âge
w_x – vecteur de pondération (vecteur de structure)

et avec l'expression vectorielle : $TB = \vec{\mathrm{T}}_x \cdot \vec{W}_x$

On voit qu'un taux brut est le produit scalaire de deux vecteurs de dimension « x »

9

Application : analyse comparative de la mortalité

Données disponibles :

- Nombre de décès
- Effectif de la population

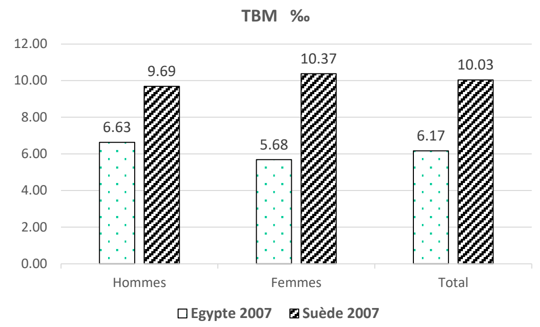
Procédure :

1. Calculer les taux de mortalité
2. Présenter les taux (bruts) graphiquement

Résultats (conclusion) :

- Mortalité en Suède est plus élevée qu'en Égypte **(vrai ou faux ?)**
- En Égypte la mortalité des hommes est plus élevée que la mortalité des femmes **(vrai ou faux ?)**
- En Suède la mortalité des femmes est plus élevée que la mortalité des hommes **(vrai ou faux ?)**

	Hommes	Femmes	Total
Décès			
Egypte 2007	248 766	204 065	452 831
Suède 2007	44 025	47 795	91 820
Population			
Egypte 2007	37 538 734	35 896 907	73 435 641
Suède 2007	4 543 729	4 607 261	9 150 990
TBM ‰			
Egypte 2007	6.63	5.68	6.17
Suède 2007	9.69	10.37	10.03



10

Analyse plus approfondie de la mortalité et du contexte démographique

Si les données (décès, population exposée) sont classées par groupes d'âge quinquennaux à l'occurrence on peut calculer et comparer :

- les taux de mortalités par âge (analyse plus approfondie)
- les structures des populations par âge et par sexe

Age	Égypte 2007						Suède 2007					
	Décès			Population			Décès			Population		
	Hommes	Femmes	Total	Hommes	Femmes	Total	Hommes	Femmes	Total	Hommes	Femmes	Total
0	21 332	18 350	39 682	1 055 864	1 009 633	2 065 497	149	121	270	55 056	51 962	107 018
1-4	5 512	4 915	10 427	3 294 926	3 136 568	6 431 494	**29**	**39**	68	211 177	203 636	414 813
5-9	2 871	1 893	4 764	4 878 002	4 572 625	9 450 627	**20**	**21**	41	242 511	229 619	472 130
10-14	2 451	1 515	3 966	5 047 224	4 697 621	9 744 845	**29**	**16**	45	283 857	270 730	554 587
15-19	3 938	1 934	5 872	4 466 896	4 096 756	8 563 652	141	**63**	204	322 963	305 413	628 376
20-24	4 543	2 330	6 873	3 264 892	2 991 590	6 256 482	196	**65**	261	281 434	268 118	549 552
25-29	3 954	2 142	6 096	2 588 705	2 814 192	5 402 897	212	**65**	277	281 427	269 667	551 094
30-34	3 734	2 289	6 023	2 473 133	2 492 762	4 965 895	193	92	285	304 553	293 071	597 624
35-39	4 461	2 932	7 393	2 362 720	2 388 210	4 750 930	261	166	427	319 908	309 062	628 970
40-44	7 840	4 318	12 158	1 986 612	1 946 670	3 933 282	445	262	707	339 489	324 486	663 975
45-49	12 924	6 304	19 228	1 723 041	1 586 419	3 309 460	642	415	1 057	297 740	288 797	586 537
50-54	19 478	10 865	30 343	1 236 096	1 262 487	2 498 583	1 081	660	1 741	294 005	288 044	582 049
55-59	24 118	14 381	38 499	974 469	865 003	1 839 472	1 743	1 119	2 862	303 607	301 953	605 560
60-64	24 425	17 214	41 639	869 726	857 419	1 727 145	2 934	1 801	4 735	308 132	305 644	613 776
65-69	25 009	19 442	44 451	628 609	524 472	1 153 081	3 372	2 226	5 598	218 493	225 028	443 521
70-74	26 824	24 754	51 578	388 434	368 006	756 440	4 152	2 907	7 059	165 129	186 186	351 315
75-79	24 122	24 461	48 583	54 450	52 102	106 552	6 111	5 029	11 140	135 655	173 828	309 483
80-84	17 696	21 667	39 363	37 439	35 825	73 264	8 405	8 313	16 718	101 424	150 817	252 241
85+	13 534	22 359	35 893	207 496	198 547	406 043	13 910	24 415	38 325	77 169	161 200	238 369
total	248 766	204 065	452 831	37 538 734	35 896 907	73 435 641	44 025	47 795	91 820	4 543 729	4 607 261	9 150 990

Il faut tout simplement diviser les décès par la population ligne par ligne pour chaque sexe et les deux sexes confondus

Les résultats de calculs sont difficiles à interpréter à l'œil

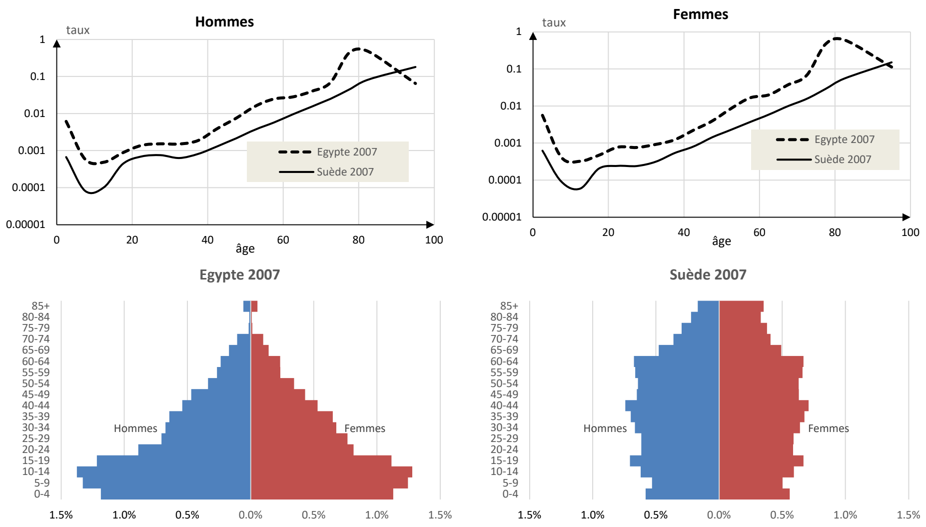
On voit toutefois qu'en Egypte les taux de mortalité par âge sont plus élevés qu'en Suède et la population suédoise est plus âgée

Age	Egypte 2007						Suède 2007					
	taux de décès			structures de la population			taux de décès			structures de la population		
	Hommes	Femmes	Total	Hommes	Femmes	Total	Hommes	Femmes	Total	Hommes	Femmes	Total
0	0.0202034	0.0202034	0.0202034	1%	1%	**3%**	0.0027063	0.0202034	0.0202034	1%	1%	**1%**
1-4	0.0016729	0.0016729	0.0016729	4%	4%	**9%**	0.0001373	0.0016729	0.0016729	2%	2%	**5%**
5-9	0.0005886	0.0005886	0.0005886	7%	6%	**13%**	0.0000825	0.0005886	0.0005886	3%	3%	**5%**
10-14	0.0004856	0.0004856	0.0004856	7%	6%	**13%**	0.0001022	0.0004856	0.0004856	3%	3%	**6%**
15-19	0.0008816	0.0008816	0.0008816	6%	6%	**12%**	0.0004366	0.0008816	0.0008816	4%	3%	**7%**
20-24	0.0013915	0.0013915	0.0013915	4%	4%	9%	0.0006964	0.0013915	0.0013915	3%	3%	**6%**
25-29	0.0015274	0.0015274	0.0015274	4%	4%	7%	0.0007533	0.0015274	0.0015274	3%	3%	**6%**
30-34	0.0015098	0.0015098	0.0015098	3%	3%	7%	0.0006337	0.0015098	0.0015098	3%	3%	**7%**
35-39	0.0018881	0.0018881	0.0018881	3%	3%	6%	0.0008159	0.0018881	0.0018881	3%	3%	**7%**
40-44	0.0039464	0.0039464	0.0039464	3%	3%	5%	0.0013108	0.0039464	0.0039464	4%	4%	7%
45-49	0.0075007	0.0075007	0.0075007	2%	2%	5%	0.0021562	0.0075007	0.0075007	3%	3%	6%
50-54	0.0157577	0.0157577	0.0157577	2%	2%	3%	0.0036768	0.0157577	0.0157577	3%	3%	6%
55-59	0.0247499	0.0247499	0.0247499	1%	1%	3%	0.0057410	0.0247499	0.0247499	3%	3%	7%
60-64	0.0280836	0.0280836	0.0280836	1%	1%	2%	0.0095219	0.0280836	0.0280836	3%	3%	7%
65-69	0.0397847	0.0397847	0.0397847	1%	1%	2%	0.0154330	0.0397847	0.0397847	2%	2%	5%
70-74	0.0690568	0.0690568	0.0690568	1%	1%	1%	0.0251440	0.0690568	0.0690568	2%	2%	4%
75-79	0.4430119	0.4430119	0.4430119	0%	0%	0%	0.0450481	0.4430119	0.4430119	1%	2%	3%
80-84	0.4726622	0.4726622	0.4726622	0%	0%	0%	0.0828699	0.4726622	0.4726622	1%	2%	3%
85+	0.0652254	0.0652254	0.0652254	0%	0%	1%	0.1802537	0.0652254	0.0652254	1%	2%	3%
total	0.0066269	0.0066269	0.0066269	51%	49%	100%	0.0096892	0.0066269	0.0066269	50%	50%	100%

On remarque que l' âge médian en Egypte = 20 ans ; en Suède = 40 ans ➔ la visualisation des données et des résultats de calculs permettra de repérer mieux la différence

La présentation graphique simplifie l'interprétation

Les conclusions sont évidentes, cependant les graphiques ne permettent pas **de quantifier la mortalité différentielle dans les deux pays** : il faut donc chercher les moyens de quantification !



13

Quatre raisons pour prendre en considération les compositions par âge

1. **Les taux de mortalité** varient beaucoup avec âge.
2. **La composition par âge des populations** varie considérablement
   - ✓ Effet associé à la nation (populations nationales)
   - ✓ Effet associé au sexe
   - ✓ Effet associé à l'état matrimonial (célibataires, marié/es, divorcé/es, veufs/ves)
   - ✓ Effet ethnoculturel
   - ✓ Effet associé au statut social (catégorie socio-professionnelle)
3. La composition de la population par âge est une caractéristique démographique important **déterminée par *l'histoire*** de la natalité, de la mortalité et de la migration ainsi que par la situation sociale et économique.
4. Données sur les décès et l'effectif de la population par âge sont le plus souvent disponibles.

14

Méthodes de comparaison des taux (bruts) : standardisation

Puisque le score des taux bruts dépend de la probabilité de décès spécifique à l'âge aussi bien que de la structure de la population par âge, il est souvent difficile d'interpréter la différence entre les taux bruts, ainsi que leur dynamique. Pour surmonter cette difficulté et rendre les taux bruts comparables on fait le recours à une procédure particulière, qu'on appelle *la standardisation* ou le calcul des *taux « comparatifs »*.

Lecture recommandée :

- Guillaume Wunsch – « Variables de confusion et indices résumés » Dans: Caselli, G., J.Vallin et G.Wunsch *Démographie: analyse et synthèse.* Vol.1: La dynamique de la population. Edition de l'INED, Paris, 2001, p.329-448
- Leridon, H. et L.Toulemon – *Démographie. Approche statistique et dynamique des populations.* Economica, Paris, 1997. (Chapitre 11: « Standardisation »), p.191-209.
- Henry, L. – *Démographie. Analyse et modèles*. Edition de l'INED, Paris, 1984, p.157-158 et p.170.

3. Méthodes de calculs des taux « comparatifs », ou la standardisation des taux bruts

- **Standardisation directe (W.Ogle, 1883)[1]**
- **Standardisation indirecte (F. Neison, 1844 / W. Farr, 1855)[2]**

Lectures pédagogique :

G.Wunsch « Variables de confusion, standardisation et indices résumés » dans G.Caselli, J.Vallin, G.Wunsch (dir) *Démographie: analyse et synthèse*. Vol. 1. *La dynamique des populations*. Ined, Paris, 2001, p.329-348

Wunsch, Guillaume J. et Évelyne Thiltgès (1995) – « Une confusion standardisée: variables confondantes et standardisation » *Genus*, vol.50, n°3-4, p.27-59

Abram J. Jaffe – *Handbook of statistical methods for demographers; selected problems in the analysis of census data*. Washington, U.S. Govt. Print. Off., 1951, chap.3 "Selected statistical methods for the standardization of population", p.43-58

Bibliographie historique :

[1)] W.Ogle – Annual Summary of Births, Deaths, and Causes of Deaths in London and other great towns 1883, 1884, p.III

[2)] F.G.P. Neison – "On a method recently proposed for conducting inquiries into the comparative sanitary condition of various districts." *Journal of the Statistical Society of London*. VII, 1844; appliqué par William Farr in "The 16th Annual Report of the Registrar-General of Births, Deaths and Marriage in England and Wales" (1855)

Voir aussi : http://isi.cbs.nl/glossary/term953.htm

J.Körösi – "Mortalitäts-Coefficient und Mortalitäts-Index" *Bulletin de l'institut International de Statistique*, t.VI, 1892

L. v. Bortkewitsch – Kritische Betrachtungen zur theoretischen Statistik*, Jahrbücher für Nationalökonomie und Statistik*, III. Folge. Jena 1896, Artikel: „Über die Methode der „Standard Population" // *Bulletin de l'Institut International de Statistique*, t.XIV, 2 liv. Berlin 1904, p.417-437, t.XI, 1 liv.1, p.173-176,178.

La standardisation « directe »

Idée de base :

Calculer pour une population (dite « le standard », ou une « **population-type** ») les taux bruts correspondants au chaque assortiment des taux de mortalité par âge observés dans des populations réelles.

On combine donc les taux par âge de chaque population avec une structure par âge type (standard).

On appelle cette méthode « standardisation directe » puisqu'elle permet de calculer directement les taux bruts comparatifs (on dit « taux comparatifs » tout court)

Données indispensables :

1. Le nombre d'événements classés ***par catégories de structure*** dans toutes les populations à comparer.
2. L'effectif de chaque catégorie de structure dans toutes les populations à comparer

Application : analyse comparative de la mortalité à réviser

Données disponibles :

- Nombre de décès
- Effectif de la population

Procédure habituelle :

1. Calculer les taux de mortalité
2. Présenter les taux (bruts) graphiquement

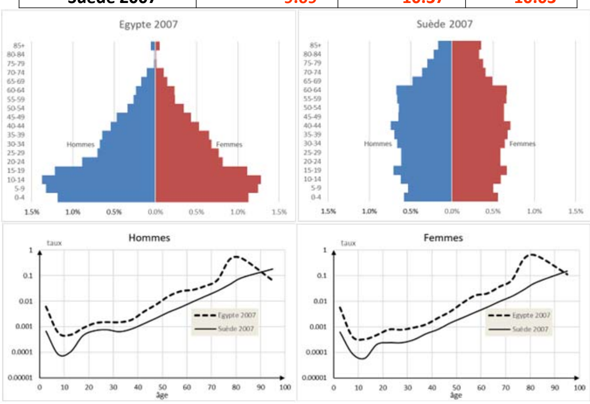
Résultats (conclusion) ne sont pas convaincants :

- Mortalité en Suède est plus élevée qu'en Égypte **(vrai ou faux ?)**
- En Égypte la mortalité des hommes est plus élevée que la mortalité des femmes **(vrai ou faux ?)**
  mais le TBD pour les hommes est plus élevé
- En Suède la mortalité des femmes est plus élevée que la mortalité des hommes **(vrai ou faux ?)**

On constate par ailleurs :

- Une différence des structures par âge
- Une différence systématique de la mortalité par âge en faveur de l'Égypte

	Hommes	Femmes	Total
	Décès		
Egypte 2007	248 766	204 065	452 831
Suède 2007	44 025	47 795	91 820
	Population		
Egypte 2007	37 538 734	35 896 907	73 435 641
Suède 2007	4 543 729	4 607 261	9 150 990
	TBM ‰		
Egypte 2007	6.63	5.68	6.17
Suède 2007	9.69	10.37	10.03



Un exercice de la standardisation directe (calculs)

Données

Age	Suède, femmes, 2007 Population[1)]	Suède, femmes, 2007 Décès	Égypte, femmes, 2007 Population[1)]	Égypte, femmes, 2007 Décès
0	51 962	121	174 078	3 720
1-4	203 636	39	754 758	1 220
5-9	229 619	21	879 129	396
10-14	270 730	16	808 510	298
15-19	305 413	63	720 161	561
20-24	268 118	65	622 988	673
25-29	269 667	65	733 057	752
30-34	293 071	92	732 312	956
35-39	309 062	166	612 825	1 113
40-44	324 486	262	487 996	1 405
45-49	288 797	415	284 799	1 226
50-54	288 044	660	503 608	2 878
55-59	301 953	1 119	301 879	3 266
60-64	305 644	1 801	374 317	5 212
65-69	225 028	2 226	256 247	6 866
70-74	186 186	2 907	154 623	6 182
75-79	173 828	5 029	149 917	8 199
70-84	150 817	8 313	88 716	9 013
85 +	161 200	24 415	58 940	10 627
Total	4 607 261	47 795	8 698 860	64 563
TBM ‰	xx	10.37	xx	7.42

1) Population au 30 juin équivaut la population exposée

Calculs des taux bruts et des taux comparatifs

	ASMR Suède $\vec{A}$	ASMR Égypte $\vec{B}$	Structure par âge Suède $\vec{C}$	Structure par âge Égypte $\vec{D}$	Structure par âge moyenne $\vec{E}$
	0.00233	0.01817	1.1%	2.8%	2.0%
	0.00019	0.00157	4.4%	8.7%	6.6%
	0.00009	0.00041	5.0%	12.7%	8.9%
	0.00006	0.00032	5.9%	13.1%	9.5%
	0.00021	0.00047	6.6%	11.4%	9.0%
	0.00024	0.00078	5.8%	8.3%	7.1%
	0.00024	0.00076	5.9%	7.8%	6.8%
	0.00031	0.00092	6.4%	6.9%	6.7%
	0.00054	0.00123	6.7%	6.7%	6.7%
	0.00081	0.00222	7.0%	5.4%	6.2%
	0.00144	0.00397	6.3%	4.4%	5.3%
	0.00229	0.00861	6.3%	3.5%	4.9%
	0.00371	0.01663	6.6%	2.4%	4.5%
	0.00589	0.02008	6.6%	2.4%	4.5%
	0.00989	0.03707	4.9%	1.5%	3.2%
	0.01561	0.06727	4.0%	1.0%	2.5%
	0.02893	0.46948	3.8%	0.1%	2.0%
	0.05512	0.60480	3.3%	0.1%	1.7%
	0.15146	0.11261	3.5%	0.6%	2.0%
somme	XX	XX	100%	100.0%	100.0%

	Suède	Égypte		
TBM ‰	10.37	5.68	$\vec{A}\bullet\vec{C}$	$\vec{B}\bullet\vec{D}$
TMC ‰ (std=C)	**10.37**	49.92	$\vec{A}\bullet\vec{C}$	$\vec{B}\bullet\vec{C}$
TMC ‰ (std=D)	1.88	**5.68**	$\vec{A}\bullet\vec{D}$	$\vec{B}\bullet\vec{D}$
TMC ‰ (std=E)	6.13	27.80	$\vec{A}\bullet\vec{E}$	$\vec{B}\bullet\vec{E}$

Taux sont calculés directement comme le produit de deux vecteurs

Pour calculer le produit scalaire de deux vecteurs avec MS Excel on applique un opérateur :
=SOMMEPROD(vecteur1; vecteur2;…)
Tous les vecteurs doivent être de même dimension (même nombre d'éléments)

Application revisitée : analyse comparative de la mortalité

Données disponibles :

- Nombre de décès
- Effectif de la population

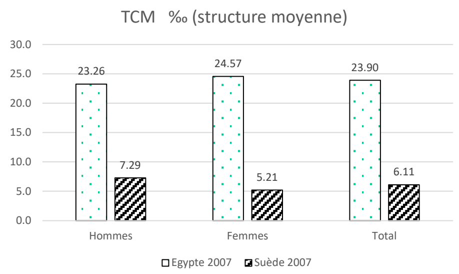
Procédure :

1. Calculer les taux de mortalité
2. Présenter les taux (bruts) graphiquement

Résultats (conclusion) :

- Mortalité en Égypte (24‰) est quatre fois plus élevée qu'en Suède (6‰)
  mais le TBM est plus élevé en Suède
- En Égypte la mortalité des hommes (23‰) est moins élevée que la mortalité des femmes (25‰)
  mais le TBM masculine est plus élevé
- En Suède la mortalité des femmes (5.2‰) est moins élevée que la mortalité des hommes (7.3‰)
  mais le TBM masculine est plus élevé

	Hommes	Femmes	Total
Décès			
Egypte 2007	**248 766**	**204 065**	**452 831**
Suède 2007	**44 025**	**47 795**	**91 820**
Population			
Egypte 2007	**37 538 734**	**35 896 907**	**73 435 641**
Suède 2007	**4 543 729**	**4 607 261**	**9 150 990**
TBM ‰			
Egypte 2007	**6.63**	**5.68**	**6.17**
Suède 2007	**9.69**	**10.37**	**10.03**
TCM ‰ (standard=population moyenne deux sexes confondus)			
Egypte 2007	**23.26**	**24.57**	**23.90**
Suède 2007	**7.29**	**5.21**	**6.11**



Facteur (ou multiplicateur) comparatif de la mortalité = changement du système métrique (cf. « Indice de Laspeyres »)

Soit $TCM = \frac{\sum_x m_x \cdot P_x^s}{\sum_x P_x^s}$ un **T**aux **C**omparatif de **M**ortalité d'une population avec les taux de mortalité par âge $\boldsymbol{m_x}$ et la population standard $\boldsymbol{P_x^S}$

en multipliant le numérateur et le dénominateur par $\sum_x m_x^s \cdot P_x^s$ (le nombre de décès dans le standard)

on obtient $TCM = \frac{\sum_x m_x \cdot P_x^s}{\sum_x P_x^s} \cdot \frac{\sum_x m_x^s \cdot P_x^s}{\sum_x m_x^s \cdot P_x^s}$

qu'on peut transformer par la permutation des termes $TCM = \frac{\sum_x m_x^s \cdot P_x^s}{\sum_x P_x^s} \cdot \frac{\sum_x m_x \cdot P_x^s}{\sum_x m_x^s \cdot P_x^s}$

où $\frac{\sum_x m_x^s \cdot P_x^s}{\sum_x P_x^s}$ est le taux brut de mortalité dans la population standard et

$FCM = \frac{\sum_x m_x \cdot P_x^s}{\sum_x m_x^s \cdot P_x^s}$ est le facteur ou le multiplicateur comparatif de mortalité (*comparative mortality figure or comparative mortality factor*) marquant la distance relative entre le niveau de mortalité dans la population standard et celui dans la population étudiée.

On peut démontrer que $FCM = \frac{\sum_x m_x^s \cdot P_x^s \cdot \left(\frac{m_x}{m_x^s}\right)}{\sum_x m_x^s \cdot P_x^s} = \frac{\sum_x d_x^s \cdot \left(\frac{m_x}{m_x^s}\right)}{\sum_x d_x^s}$ est la moyenne de ratios de *taux de mortalité par âge* pondérés par la distribution des décès par âge de la population standard

Un exercice de la standardisation directe en résumé

1. Calculons des taux comparatifs (bruts standardisés) comme si la structure de la population en Égypte était la même qu'en Suède, alors on pourrait attribuer la différence entre les taux bruts de mortalité de deux pays à la différence de la mortalité par âge :

$$TCM_{Suède}^{Suède} = \sum_x {}_n m_x^{Egypte} \cdot {}_n C_x^{Suède} \quad et \quad TCM_{Egypte}^{Suède} = \sum_x {}_n m_x^{Egypte} \cdot {}_n C_x^{Suède}$$

$$TCM_{Suède}^{Suède} = 10{,}03‰ \text{ et } TCM_{Egypte}^{Suède} = 42{,}11‰$$

2. On peut inverser la situation et calculer les TBM pour la Suède comme si elle a la même structure de la population que le Kazakhstan, les valeurs des « taux comparatifs » ne sont pas les même que dans le premier exercice, mais le rapporte entre eux reste à peu près le même :

$$TCM_{Suède}^{Egypte} = \sum_x {}_n m_x^{Suède} \cdot {}_n C_x^{Egypte} \quad et \quad TCM_{Egypte}^{Egypte} = \sum_x {}_n m_x^{Egypte} \cdot {}_n C_x^{Egypte}$$

$$TCM_{Suède}^{Egypte} = 2{,}25‰ \quad et \quad TCM_{Egypte}^{Egypte} = 6{,}17‰$$

Le choix d'un standard

Standard vieux :

$$TCM_{Suède}^{Suède} = 10,03‰$$

$$TCM_{Egypte}^{Suède} = 42,11‰$$

Standard jeune :

$$TCM_{Suède}^{Egypte} = 2,25‰$$

$$TCM_{Egypte}^{Egypte} = 6,17‰$$

Solution 1 : standard « moyen »

Pour comparer la population A avec la population B le standard est la structure moyenne

$$C_x^s = \frac{C_x^A + C_x^B}{2}$$

Standard moyen:

$$\boldsymbol{TBM^m}_{\boldsymbol{Suède}} = \mathbf{6,14‰}$$

$$\boldsymbol{TBM^m}_{\boldsymbol{Egypte}} = \mathbf{23,50‰}$$

Solution 2 : population type → une variante d'une population « moyenne » (ex. standard européen, voir le fichier sur EPI)

Approche intéressante : le standard « population moyenne » est une combinaison linéaire des toute les populations comparées s'interprète comme composant principal dans les termes d'analyse factorielle

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

23

23

La standardisation « indirecte »

Idée de base :

Estimer la différence (rapport) entre le nombre de décès observé dans une population, avec celui qui pourrait avoir lieu, si la mortalité par âge était comme dans une population de référence (mortalité standard).

Ainsi on peut calculer une série des taux comparatifs (« standardisés ») par rapport à un taux brut d'une population de référence.

Donc on combine les structures par âge de chaque population avec **un même assortiment des taux par âge (taux-type).**

On appelle cette méthode « standardisation indirecte » puisque on compare les niveaux de la mortalité par intermédiaire d'un indicateur médiat.

Données indispensables:

- ✓ Le nombre total d'événements pour toutes les populations à comparer.
- ✓ Le nombre d'événements classés par catégories de structure au moins dans une des populations à comparer.
- ✓ L'effectif de chaque catégorie de structure dans toutes les populations à comparer.

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

24

24

Rapport comparatif de mortalité *(fr)*
Comparative Mortality Ratio (en)
(cf. « Indice de Paasche »)

Le rapport entre le nombre observé de décès dans la population et le nombre espéré de décès sous la condition que la mortalité par âge correspond à un standard.

Si dans la population (A) la distribution de décès par âge est inconnue on peut quand même mesurer le niveau général de sa mortalité par rapport au niveau standard (population B en l'occurrence) comme suit :

$$RCM = \frac{\sum_x m_x^A \cdot P_x^A}{\sum_x m_x^B \cdot P_x^A} = \frac{D^A}{\sum_x m_x^B \cdot P_x^A} = \frac{Nombre\ observés\ de\ décès}{Nombre\ espéré\ de\ décès}$$

où

$\boldsymbol{D^A}$ **=> nombre de décès tous âges confondus dans la population A ;**

$\boldsymbol{P_x^A}$ **=> effectif du groupe dans l'intervalle d'âge « x » dans la population A**

$\boldsymbol{m_x^B}$ **=> taux de mortalité dans l'intervalle d'âge « x » dans la population B (standard)**

Nota : le RCM est un frère jumeau du FCM $FCM = \frac{\sum_x m_x \cdot P_x^s}{\sum_x m_x^s \cdot P_x^s}$

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

25

25

Exercice de la standardisation indirecte (calculs)

Données

Age	Suède, 2007		Égypte, 2007	
	Population[1)]	Décès	Population[1)]	Décès
	$\vec{A}$		$\vec{B}$	
0	107 018	~~270~~	2 065 497	39 682
1-4	414 813	~~68~~	6 431 494	10 427
5-9	472 130	~~41~~	9 450 627	4 764
10-14	554 587	~~45~~	9 744 845	3 966
15-19	628 376	~~204~~	8 563 652	5 872
20-24	549 552	~~261~~	6 256 482	6 873
25-29	551 094	~~277~~	5 402 897	6 096
30-34	597 624	~~285~~	4 965 895	6 023
35-39	628 970	~~427~~	4 750 930	7 393
40-44	663 975	~~707~~	3 933 282	12 158
45-49	586 537	~~1 057~~	3 309 460	19 228
50-54	582 049	~~1 741~~	2 498 583	30 343
55-59	605 560	~~2 862~~	1 839 472	38 499
60-64	613 776	~~4 735~~	1 727 145	41 639
65-69	443 521	~~5 598~~	1 153 081	44 451
70-74	351 315	~~7 059~~	756 440	51 578
75-79	309 483	~~11 140~~	106 552	48 583
70-84	252 241	~~16 718~~	73 264	39 363
85 +	238 369	~~38 325~~	406 043	35 893
Total	9 150 990	91 820	73 435 641	452 831
TBM ‰	xx	10.03	xx	6.17
Décès attendu	$\vec{A}$ x $\vec{D}$= **385 309**			xx
RCM	91 820 : 385 309=0.24			1
TMC	0.24 x 6.17 = **1.47**			6.17

1) Population au 30 juin équivaut la population exposée

Calculs

ASMR*/TMSA**		
Suède	Égypte	moyenne
$\vec{C}$	$\vec{D}$	$\vec{E}$
0.00252	0.01921	0.01087
0.00016	0.00162	0.00089
0.00009	0.00050	0.00030
0.00008	0.00041	0.00024
0.00032	0.00069	0.00051
0.00047	0.00110	0.00079
0.00050	0.00113	0.00082
0.00048	0.00121	0.00084
0.00068	0.00156	0.00112
0.00106	0.00309	0.00208
0.00180	0.00581	0.00381
0.00299	0.01214	0.00757
0.00473	0.02093	0.01283
0.00771	0.02411	0.01591
0.01262	0.03855	0.02559
0.02009	0.06819	0.04414
0.03600	0.45596	0.24598
0.06628	0.53728	0.30178
0.16078	0.08840	0.12459

* - age specific mortality rate (en)
** - taux de mortalité spécifique à l'âge (fr)

Le taux comparatif se fabrique **indirectement** (en passant par le calcul du RCM)

1° On applique cette méthode, si la répartition des décès (${}_nD_x$) par âge en Suède n'est pas disponible, mais on connait la répartition de la population par âge (${}_nP_x$) et le nombre total de décès ($D = \sum {}_nD_x$) en Suède.

2° On calcule les taux de mortalité par âge pour l'Égypte :

$${}_nm_x^E = \frac{{}_nD_x^E}{{}_nP_x^E}$$

3° Hypothèse 0 : *si la mortalité (par âge) en Suède est la même qu'en Égypte*, l'application des taux de mortalité par âge en Égypte à la population suédoise

$$D^S = \widehat{D}^S = \sum_x {}_nm_x^E \cdot {}_nP_x^S = \sum_x \widehat{D}_x^S$$

produira le même nombre de décès qu'on y observe (en l'occurrence 91820).

Si *la mortalité en Égypte est plus élevée*, ses taux de mortalité par âge produisent plus de décès en Suède qu'on y observe :

$$D^S < \widehat{D}^S \to RCM = \frac{D^S}{\widehat{D}^S} < 1$$

Dans le cas contraire :

$$D^S > \widehat{D}^S \to RCM = \frac{D^S}{\widehat{D}^S} > 1$$

Faites les calculs avec les autres standards (Suède et une suite des « taux-type » moyens)

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

26

26

Un exercice de la standardisation indirecte en résumé

1. Si la mortalité par âge en Suède était comme en Égypte (standard=Égypte), le nombre de décès en Suède était 385 309 :
   $RMC_{Suède}$ = 91 820 : 385 309 ≈ 0,24
   et le taux brut de mortalité était (taux comparatif) ≈ 4,8 et non 10,55

$$TBM^{C}_{Suède} = RCM_{Suède} \times TBM_{Egypte} = 0,24 \cdot 6,17 = 1,47 \qquad \mathrm{TBM}_{standard} = 6,17$$

Conclusion : la mortalité en Suède est 4,2 (1/0,24) fois inférieure de la mortalité en Égypte

2. Si la mortalité par âge en Égypte était comme en Suède (standard=Suède), le nombre de décès y était 165 309 :
   RMC_{Egypte} = 452 831 : 165 309 ≈ 2,74
   et le taux brut de mortalité y était ≈ 27,49 et non 6,17

$$TBM^{C}_{Egypte} = RCM_{Egypte} \times TBM_{Suède} = 2,74 \cdot 10,03 = 27,49 \qquad \mathrm{TMB}_{Standard} = 10,03$$

Conclusion : la mortalité en Égypte est 2,7 fois supérieure de la mortalité en Suède

Algorithme général de calculs (standardisation indirecte)

$$TMC^{A} = RMC^{A} \cdot TBM^{snd} =$$

$$= \frac{D^{A}}{\sum_{x} m_x^{snd} \cdot P_x^{A}} \cdot \frac{\sum_{x} m_x^{snd} \cdot P_x^{snd}}{\sum_{x} P_x^{snd}}$$

TMC^{A} – taux de mortalité (brut) comparatif pour la population A;

P_x^{i} – effectif du groupe d'âge *x* dans une population i (i=A, i=standard)

m_x^{std} – taux de mortalité par group (d'âge) dans une population standard

D^{A}– nombre de décès dans la population A pour laquelle on calcule le taux brut comparatif

Register-General de Grande Bretagne applique cette méthode pour comparer la mortalité des groupes SP.

1[er] exemple d'applications de la standardisation indirecte:

1. Rapport comparatif de la mortalité (*standardized mortality ratio* en USA) est un indicateur des risques relatifs de décès selon les catégories socioprofessionnelles.

Age	Nombre des mineurs de sexe masculin P_x	Taux de mortalité de tuberculeuse (tous les actifs) M_x
20-24	74 589	0,0001260
25-29	85 077	0,0001612
30-34	80 845	0,0002154
35-44	148 870	0,0003396
45-54	102 649	0,0005682
55-59	42 494	0,0007526
60-64	30 037	0,0008237
Total 20-64	564 570	0,0009565

Source: Mary McGehee, « Mortality ». In J.S.Siegel , D.A.Swanson. (eds) *Methods and Materials of Demography*, 2d ed. p.281

$$SMR = \frac{d}{\sum M_x \cdot P_x} = \frac{d}{D^{\exp}} = \frac{540}{206} = 2,63$$

Nombre attendu de décès des mineurs

$$D^{exp} = \sum M_x \cdot P_x = 206$$

En 1950 *US National Office of Vital Statistics* a enregistré ***540 décès des mineurs à cause de tuberculeuse*** d'où le

$$RCM = \frac{540}{206} = 2,6$$

<u>Conclusion :</u> un mineur a ***2,6 fois plus de risque*** de décéder d'un tuberculeuse qu'un travailleur « moyen ».

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

29

29

2[d] exemple d'applications de la standardisation indirecte:

2. ***<u>Rapport proportionnel de mortalité standardisé par âge</u>***[(1)] utilisé (pour la première fois ?) en 1997 dans le rapport de *U.S. NIOSH*[(2)]*: 24 Reporting Stats, 1984-1988, by C.Burnett et al*. pour estimer les risques de décès selon les catégories socioprofessionnelles: *(4 groups sexe–race) x (3 groups d'âge) x (325 catégories d'emploi) x (235 catégories d'industrie) x (188/192 causes de décès pour femmes/hommes) = 174 135 000 combinaisons*

Catégories socioprofessionnelles	Cause X	Autres causes	Toutes causes
Catégorie Y	A_i	B_i	N_{1i}
Autres catégories	C_i	D_i	N_{2i}
Toutes catégories	M_{1i}	M_{2i}	Ti

Soit

$$E(A_i) = \frac{M_{1i} \cdot N_{1i}}{T_i} \equiv \frac{M_{1i}}{T_i} \cdot N_{1i}$$

un nombre attendu de décès pour une catégorie professionnelle (Y=1), d'une cause de décès (X) et group d'âge (i), alors

$$RPM = \frac{\sum A_i}{\sum E(A_i)} \cdot 100$$

<u>Une avantage</u> : on n'a pas besoin de données qui sont indispensables pour les autres méthodes de standardisation.

<u>Les désavantages</u> :

1. Ce n'est pas un taux mais un rapport ou un ratio (dénominateur=nombre total de décès)
2. Si le risque moyen pour une branche d'industrie est faible, RPM peut surestimer les risques relatifs, et inversement, il sous-estime les risques, là où le risque moyen est fort.
3. Le RPM pourrait aussi être biaisé par une très forte mortalité d'une cause ou des causes majeures de décès (sous-estimation des risques, si le nombre de décès d'une cause est trop élevé et inversement)
4. Le RPM peut être affecté, si des catégories non-déclarées ne sont pas distribuées proportionnellement

(1) Cf. McGehe « Mortality », in J.S.Siegel and D.A.Swanson (ed), *The Methodes and Materials of Demography*, 2d edition, Emerald, 2008, p.282

(2) National Institute for Occupational Safety & Health (https://www.cdc.gov/niosh/index.htm)

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie 30

30

Une autre idée de la comparaison des taux bruts, la standardisation « inverse »

Idée de base :

Comparer l'effectif d'une population avec celui que cette population devrait avoir pour produire tel nombre de décès, qu'elle produise, si ses taux de la mortalité par âge étaient comme dans une population de référence (mortalité standard, nombre de décès par âge connu, effectif total connu). Cette méthode est « inverse » puisque à la place de comparer directement les taux de mortalité, ou des nombres de décès on estime et compare l'effectif d'une population.

Données indispensables:

- ✓ Le nombre d'événements classés par catégories de structure pour toutes les populations à comparer.
- ✓ L'effectif de la population par âge au moins dans une des populations à comparer (ou les taux standards).
- ✓ Les effectifs totaux de toutes les populations à comparer (les structures sont inconnues).

Méthode de calculs des taux comparatifs :

$$TC = \frac{\sum_x \frac{N_x}{t_x^s}}{P} \cdot TB^{st}$$

TC– taux comparatif (taux brut standardisé)

N_x – nombre d'événements classés par âge (catégorie de structure)

t_x^s – taux spécifique à l'âge (catégorie) dans une population standard (taux-type)

P – effectif de population à étudier

TB^{st} – taux brut pour la population standard

Une application (et la découverte) de cette méthode a été proposée dans:
Calot Gérard. « Pour une estimation rapide de l'indicateur conjoncturel de la fécondité » // *Population*, 33e année, n°3, 1978. pp. 705-716;

Problème de prise en considération deux (ou plus) facteurs structuraux

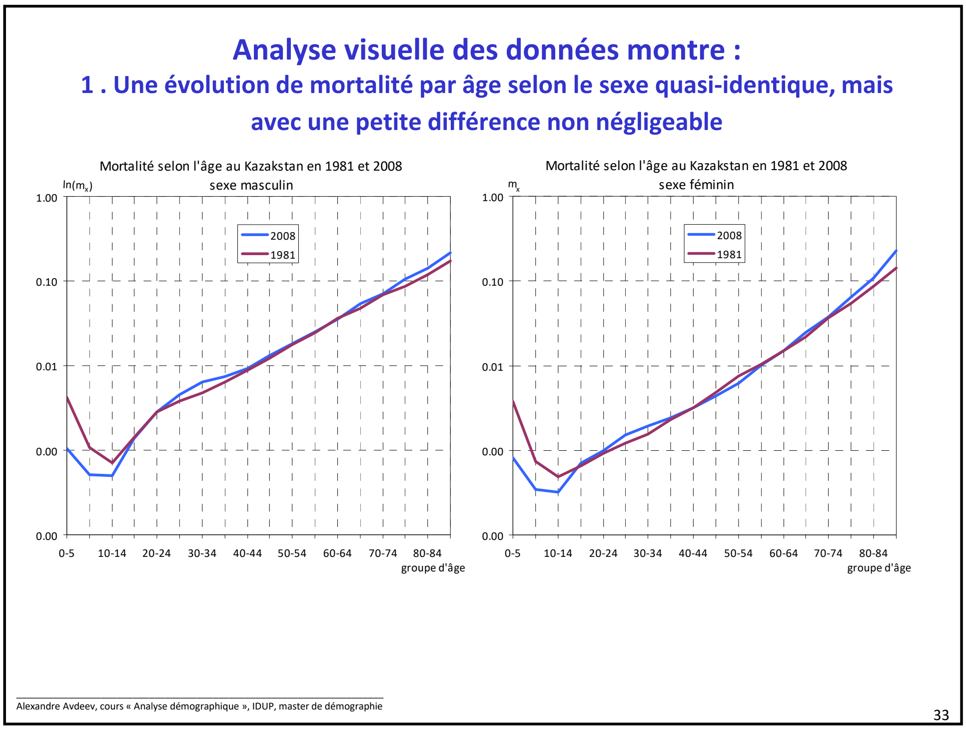
La dynamique de mortalité dépend des plusieurs dimensions de structure superposées, e.g. de celle par âge et celle par sexe.

Comparons les changements de la mortalité générale au Kazakhstan entre 1981 et 2008

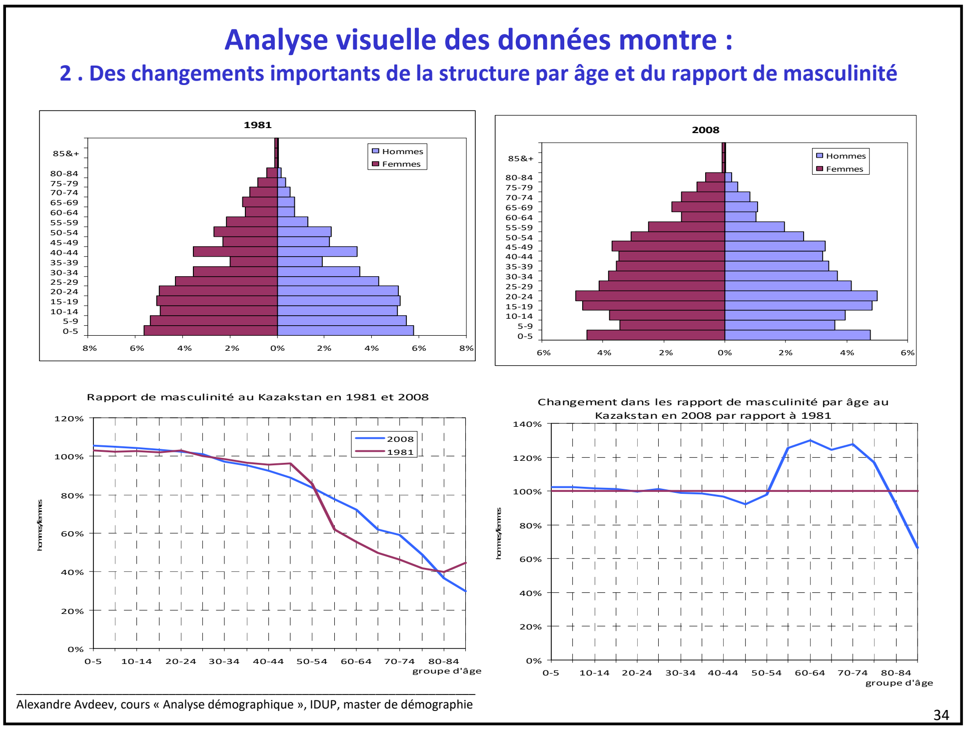
On pourrait prendre comme standard une des sous-populations (e.g. hommes en 2008) et calculer les taux comparatifs pour les autres.

Caractéristiques		TBM (observé)			TCM (standard homme 2008)		
Sexe		1981	2008	variation	1981	2008	variation
	Hommes	9.132	11.245	2.11	10.933	11.245	0.312
	Femmes	6.98	8.296	1.32	5.584	5.438	-0.146
	Les deux sexes	8.017	9.715	1.698	7.795	7.962	0.168
Rapport de masculinité (hommes pour 100 femmes)		92.9	92.7	-0.2	100	100	0

On peut accepter et interpréter ce résultat, mais de fait un tel calcul ne prend pas en considération les changements du niveau de mortalité apportés par la variation du rapport de masculinité


Analyse visuelle des données montre :
1 . Une évolution de mortalité par âge selon le sexe quasi-identique, mais avec une petite différence non négligeable
Mortalité selon l'âge au Kazakstan en 1981 et 2008
sexe masculin
Mortalité selon l'âge au Kazakstan en 1981 et 2008
sexe féminin
2008
1981
groupe d'âge
Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie
33

33


Analyse visuelle des données montre :
2 . Des changements importants de la structure par âge et du rapport de masculinité
1981
2008
Hommes
Femmes
Rapport de masculinité au Kazakstan en 1981 et 2008
Changement dans les rapport de masculinité par âge au Kazakstan en 2008 par rapport à 1981
groupe d'âge
Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie
34

34

Double standardisation : prise en considération deux (ou plus) facteurs structuraux

$$TCM = \sum m_{2008}^{h} \cdot c_{1981}^{h} + \sum m_{2008}^{f} \cdot c_{1981}^{f}$$

où $${}^{a}c_{x}^{s} = \frac{{}^{a}p_{x}^{s}}{\sum_{x}\sum_{s} {}^{a}p_{x}^{s}}$$ telle qu'on calcule pour dessiner une pyramide de la population

Donc on standardise le niveau de mortalité générale sous une hypothèse la structure de la population par âge et par sexe ne change pas.

Les résultats de calcul montrent que TCM 2008 = 8.19‰ (TBM 2000 = 9,7‰)
et en 1981 TBM =8.02‰

Conclusion : si on prend en considération tous les facteurs structuraux, on dirait que la mortalité au Kazakhstan a augmenté en 2008 par rapport à 1981, malgré une baisse importante de la mortalité infantile et juvénile

4. Décomposition d'une différence entre les taux bruts

- **Modèles additifs sans et avec interaction.**
- **Modèles multiplicatifs**


Décomposition de la différence entre deux taux bruts


Evelyn M. Kitagawa, 1920-2007


Philip M. Hauser, 27.09.1909-13.12.1994


Josiah King, The University of Edinburgh School of Philosophy, Psychology and Language Sciences

Evelyn M. Kitagawa – « Components of Difference between Two Rates » *Journal of the American Statistical Association,* vol.50, no. 272, (dec.1955), p.1168-1194 (https://doi-org.ezpaarse.univ-paris1.fr/10.2307/2281213)

Evelyn M. Kitagawa and Philip M. Hauser – *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology,* Cambridge, Mass: Harvard University Press, 1973 (ISBN 9780674188471)

Prithwis Das Gupta – *Standardization and Decomposition of Rates: A User Manual*. Current Population Reports, Special Studies, Series P23-186, US Department of Commerce, Economic and Statistics Administration, Bureau of the Census, October 1993 (https://www.census.gov/library/publications/1993/demo/p23-186.html)

Josiah King – *R package for Prithwis Das Gupta's 1991 specification of decomposition and standardisation of rates*. https://github.com/josiahpjking/DasGuptR

Soit

P – une population (P=A,B, etc.)

i – une catégorie de structure de ces populations telle que $\sum_i P_i = P$

$\boldsymbol{m_i^P}$ – taux de mortalité dans la catégorie ***i*** de la population ***P***;

$\boldsymbol{C_i^P}$ – part de la population de la catégorie ***i*** dans le total de la population ***P*** telle que $\sum_i C_i^P = 1$

Δ – différence entre les taux bruts de mortalité (TBM) des populations A et B

$$\Delta = TBM^B - TBM^A = \sum_i m_i^B \cdot C_i^B - \sum_i m_i^A \cdot C_i^A$$

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie 37

37

Décomposition de la différence entre les taux bruts

(exposé méthodologique d'après S. Preston et al.,2001, p.28-29)

$$\Delta = TBM^B - TBM^A = \sum_i m_i^B \cdot C_i^B - \sum_i m_i^A \cdot C_i^A \qquad (1)$$

On peut introduire dans l'équation (1) les taux standardisés par la structure de la population A et par la structure de la population B (méthode de standardisation directe), une moitié de chaque avec signe « + » et une autre moitié avec signe « – »; et pour que la nouvelle équation reste homogène divisons chaque élément de l'équation 1 en deux parties égales comme le suit :

$$\Delta = \boxed{\frac{\sum_i m_i^B \cdot C_i^B}{2} + \frac{\sum_i m_i^B \cdot C_i^B}{2}} - \boxed{\frac{\sum_i m_i^A \cdot C_i^A}{2} - \frac{\sum_i m_i^A \cdot C_i^A}{2} +}$$

$$\boxed{+\frac{\sum_i m_i^A \cdot C_i^B}{2} - \frac{\sum_i m_i^A \cdot C_i^B}{2}} \boxed{+\frac{\sum_i m_i^B \cdot C_i^A}{2} - \frac{\sum_i m_i^B \cdot C_i^A}{2}}$$

On peut regrouper les huit membres de cette équation en quatre groupes (comme c'est indiqué par des flèches)
et ensuite – en deux …

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie 38

38

suite:

Standardisation directe par structure B	Standardisation directe par structure A	Standardisation indirecte par mortalité B	Standardisation indirecte par mortalité A
I	II	III	IV

$$\Delta = \frac{\sum_i m_i^B \cdot C_i^B}{2} - \frac{\sum_i m_i^A \cdot C_i^A}{2} + \frac{\sum_i m_i^B \cdot C_i^B}{2} - \frac{\sum_i m_i^A \cdot C_i^A}{2} +$$

$$+ \frac{\sum_i m_i^A \cdot C_i^B}{2} - \frac{\sum_i m_i^B \cdot C_i^A}{2} + \frac{\sum_i m_i^B \cdot C_i^A}{2} - \frac{\sum_i m_i^A \cdot C_i^B}{2}$$

$$\Delta = \left[\sum_i C_i^B \left(\frac{m_i^B + m_i^A}{2}\right) - \sum_i C_i^A \left(\frac{m_i^B + m_i^A}{2}\right)\right] +$$

$$+ \left[\sum_i m_i^B \left(\frac{C_i^B + C_i^A}{2}\right) - \sum_i m_i^A \left(\frac{C_i^B + C_i^A}{2}\right)\right]$$

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

39

39

Formule définitive de la décomposition de l'effet total

$$\Delta = \sum_i \left(C_i^B - C_i^A\right) \cdot \left(\frac{m_i^B + m_i^A}{2}\right) + \sum_i \left(m_i^B - m_i^A\right) \cdot \left(\frac{C_i^B + C_i^A}{2}\right)$$

La différence des structures par âge x **Pondérée par la mortalité moyenne par âge** + **La différence des mortalités par âge** x **Pondérée par la structure moyenne par âge**

$\Delta =$ **Changement imputable à la composition de la population par âge** ***(effet de composition)*** + **Changement imputable à la différence des taux de mortalité par âge** ***(effet de taux)***

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

40

40

Exemple numérique de décomposition

Taux brut de mortalité en 2007 :

Égypte . . . . 6,17‰
Suède . . . . 10,03‰

Égypte – Suède= 6,17‰ – 10,03‰ = – 3,87‰

Cette différence (Δ = – 3,87‰) peut être décomposée en :

a) l'effet de la différence des structures de la population par âge = **– 21,86‰** (au profit de l'Égypte) ;

b) l'effet de la différence du niveau de la mortalité par âge = **+ 17,99‰** (au profit de la Suède)

→ Δ=a+b = 17,99‰ – 21,86‰ = 3,87‰

Écriture alternative pour l'analyse de variance (ANOVA) :

Le niveau (S) = le niveau (E) + l'effet de la structure + l'effet du niveau

10,03‰ = 6,17 ‰ – 21,86 ‰ + 17,99 ‰

Décomposition avec la prise en considération de l'interférence (« l'interaction »)

Soit Δ la différence entre deux taux bruts :

$$\Delta = TBM^B - TBM^A = \sum_i m_i^B \cdot C_i^B - \sum_i m_i^A \cdot C_i^A \qquad (1)$$

avec i – catégorie de structure ; m_i^p - taux spécifique à la catégorie i dans la population p ; et C_i^p - proportion de la catégorie i dans la population p

et la différence des taux $\Delta_i^m = m_i^B - m_i^A$ alors $m_i^A = m_i^B + \Delta_i^m$

de même

$$\Delta_i^C = C_i^B - C_i^A \quad \rightarrow \quad C_i^A = C_i^B + \Delta_i^C$$

On peut donc réécrira l'équation (1) en y introduisant Δ_i^C et Δ_i^m

suite:

$$\Delta = \sum_i m_i^B \cdot C_i^B - \sum_i m_i^A \cdot C_i^A =$$

$$= \sum_i \left[\left(m_i^A + \Delta_i^m \right) \cdot \left(C_i^A + \Delta_i^C \right) \right] - \sum_i m_i^A \cdot C_i^A =$$

$$= \sum_i \left(\underline{m_i^A \cdot C_i^A} + m_i^A \cdot \Delta_i^C + \Delta_i^m \cdot C_i^A + \Delta_i^m \cdot \Delta_i^C \right) - \sum_i \underline{m_i^A \cdot C_i^A} =$$

$$= \sum_i m_i^A \cdot \Delta_i^C + \sum_i \Delta_i^m \cdot C_i^A + \sum_i \Delta_i^m \cdot \Delta_i^C$$

$$\sum_i m_i^A \cdot \left(C_i^B - C_i^A \right) + \sum_i C_i^A \cdot \left(m_i^B - m_i^A \right) + \sum_i \left(C_i^B - C_i^A \right) \cdot \left(m_i^B - m_i^A \right)$$

ou

$$= \sum_i m_i^B \cdot \left(C_i^B - C_i^A \right) + \sum_i C_i^B \cdot \left(m_i^B - m_i^A \right) - \sum_i \left(C_i^B - C_i^A \right) \cdot \left(m_i^B - m_i^A \right)$$

Les inconvénients de l'interférence ou de « l'interaction »

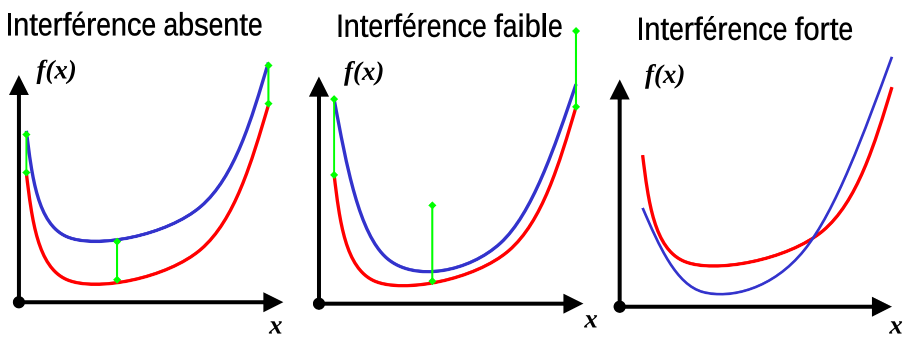
- ✓ Elle peut être positive ou négative (la covariance);
- ✓ Elle entre dans la formule avec la signe positive ou négative;

$$\Delta = \boxed{\sum_i m_i^A \cdot \left(C_i^B - C_i^A \right)} + \boxed{\sum_i C_i^A \cdot \left(m_i^B - m_i^A \right)} + \boxed{\sum_i \left(C_i^B - C_i^A \right) \cdot \left(m_i^B - m_i^A \right)}$$

Effet de la variation de la composition des populations + Effet de la variation des taux de mortalité + (–) Effet dû à la variation conjointe (« interférence ») des taux de mortalité et de la composition des populations

L'interaction (interférence) entre deux distributions:

Voir: **Wunsch, Guillaume J. et Eveline Thiltgès** (1995) – « Une confusion standardisée: variables confondantes et standardisation » *Genus,* vol.50, n°3-4, p.27-59 *



* - image modifiée par rapport à la source d'origine

Exemple numérique de décomposition avec la prise en compte de l'interférence

Taux brut de mortalité en 2007, deux sexes confondus :

Suède . . . . 10,03‰
Égypte. . . . . 6,17‰

Suède – Égypte = 10,03‰ – 6,17‰ = 3,87%

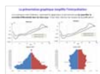
Cette différence (3,87‰) peut être décomposée en :

a) l'effet de la différence des structures de la population par âge = +7,78‰ (au profit de l'Égypte),

b) l'effet de la différence des niveaux de la mortalité = – 32,07‰ (au profit de la Suède)

c) l'effet de l'interférence structurelle = +28,16‰ (pour l'Egypte)

→ ***7,78‰ – 32,07‰ + 28,16‰ = 3,87‰***



Écriture alternative pour l'analyse de variance (ANOVA) :

$$\mu_i = \mu_s + \alpha + \beta + \alpha \cdot \beta$$

Le niveau (S) = le niveau (E) + l'effet de la structure + l'effet du niveau + l'effet de interaction des facteurs

10,03 ‰ = 6,17 ‰ + 7,78 ‰ – 32,07 ‰ + 28,16 ‰

Modèle multiplicatif de la décomposition sans interaction

Soit k – facteur multiplicateur tel que dans les populations A et B

$$TBM^B = k \cdot TBM^A$$

Par conséquent : $k = \frac{TBM^B}{TBM^A} = \frac{\sum_i m_i^B \cdot C_i^B}{\sum_i m_i^A \cdot C_i^A}$ (2)

En transformant (2) avec un artifice $k = \frac{\sum_i m_i^B \cdot C_i^B}{\sum_i m_i^A \cdot C_i^A} \times \frac{\sum_i m_i^A \cdot C_i^B}{\sum_i m_i^A \cdot C_i^B}$ et à l'aide d'une permutation

on obtient $k = \frac{\sum_i m_i^B \cdot C_i^B}{\sum_i m_i^A \cdot C_i^A} = \underbrace{\frac{\sum_i m_i^A \cdot C_i^B}{\sum_i m_i^A \cdot C_i^A}}_{\text{Indice de structure}} \times \underbrace{\frac{\sum_i m_i^B \cdot C_i^B}{\sum_i m_i^A \cdot C_i^B}}_{\text{Indice de niveau}} = I_{structure} \times I_{niveau}$

Indice de structure valeur attendue en absence de changements des taux

Indice de niveau valeur attendue, si les changements de structures sont identiques

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie 47

47

Ernst Louis Étienne Laspeyres (28.11.1834 – 4.08.1913)

Modèle multiplicatif avec l'interaction

(un analogue des indices de Laspeyres – Paasche)

Hermann Paasche (24.02.1851 – 11.04.1925)

Soit $I'_{struc} = \frac{\sum_i m_i^A \cdot C_i^B}{\sum_i m_i^A \cdot C_i^A}$ un effet de variation de la structure par âge

et $I'_{niv} = \frac{\sum_i m_i^B \cdot C_i^A}{\sum_i m_i^A \cdot C_i^A}$ un effet de variation des taux

on voit qu'après la multiplication des ce deux indices il reste un excès indécomposable que l'on peut considérer comme un effet d'interaction

$$k = \frac{TBM^B}{TBM^A} = \frac{\sum_i m_i^B \cdot C_i^B}{\sum_i m_i^A \cdot C_i^A} \neq I'_{struc} \times I'_{niv}$$

Ajoutons-y un multiplicateur pour rééquilibrer l'équation (2)

$$k = \frac{TBM^B}{TBM^A} = (I'_{struc} \times I'_{niv}) \times \frac{k}{(I'_{struc} \times I'_{niv})}$$

Interaction

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

48

48

Passage d'un modèle multiplicatif à un modèle additif

Considérons un modèle multiplicatif : $k = I_s \times I_n \Rightarrow$

On sait que

$$\ln k = \ln I_s + \ln I_n \Rightarrow 1 = \frac{\ln I_s}{\ln k} + \frac{\ln I_n}{\ln k}$$

Donc:

$$100\% = \boxed{\left(\frac{\ln I_s}{\ln k}\right) \cdot 100\%} + \boxed{\left(\frac{\ln I_n}{\ln k}\right) \cdot 100\%}$$

Effet en pourcentage imputable au changement de structure

Effet en pourcentage imputable au changement de niveau

P. Das Goupta démonstration pour deux, trois ou plus de facteurs

Soit $\boldsymbol{\alpha}$ et $\boldsymbol{\beta}$ deux facteurs tels que un taux $\boldsymbol{R}$ se fabrique comme suit : $\boldsymbol{R = \alpha\beta}$

Supposons $\boldsymbol{\alpha = A}$ et $\boldsymbol{\beta = B}$ pour la population 1 avec le taux $\boldsymbol{R_1 = AB}$
et $\boldsymbol{\alpha = a}$ et $\boldsymbol{\beta = b}$, pour la population 2 avec le taux $\boldsymbol{R_2 = ab}$

La standardisation par facteur $\boldsymbol{\beta}$ (***beta standardisation***) :

pour la population 1 → $\boldsymbol{R_1 = \frac{B+b}{2}A}$, et pour la population 2 → $\boldsymbol{R_2 = \frac{B+b}{2}a}$

La standardisation par facteur $\boldsymbol{\alpha}$ (***alfa standardisation***) :

pour la population 1 → $\boldsymbol{R_1 = \frac{A+a}{2}B}$, et pour la population 2 → $\boldsymbol{R_2 = \frac{A+a}{2}b}$

On définit donc $\boldsymbol{\alpha}$**-effet** = $\boldsymbol{\frac{B+b}{2}(A-a)}$ et $\boldsymbol{\beta}$**-effet** = $\boldsymbol{\frac{A+a}{2}(B-b)}$

Alors la différence des taux se décompose

$$\boldsymbol{R_1 - R_2 = \alpha\text{-effet} + \beta\text{-effet}}$$

De même la décomposition se fait **avec trois facteurs** $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ et $\boldsymbol{\gamma}$ ou plus $\boldsymbol{R = \alpha\beta\gamma}$

Soit $\boldsymbol{R_1 = ABC}$ et $\boldsymbol{R_2 = abc}$ on aura donc

$\boldsymbol{\alpha}$-effet = $\left[\frac{bc+BC}{3} + \frac{bC+Bc}{6}\right](A-a)$;

$\boldsymbol{\beta}$-effet = $\left[\frac{ac+AC}{3} + \frac{aC+Ac}{6}\right](B-b)$ et

$\boldsymbol{\gamma}$-effet = $\left[\frac{ab+AB}{3} + \frac{aB+Ab}{6}\right](C-c)$ etc....

Gupta, P. D. (1991). Decomposition of the difference between two rates and its consistency when more than two populations are involved. *Mathematical Population Studies*, *3*(2), 105–125.

W. Das Goupta démonstration (exemple numérique 1)

Table 2.1. Mean Earnings as the Product of Two Factors for Black Males and White Males, 18 Years and Over: United States, 1980

Measures	Black males (population 1)	White males (population 2)
Mean earnings = $\frac{Total\ earning}{Total\ population} = R$	\$ 7 846,56 (=R_1)	\$ 13 703,73 (=R_2)
Mean wage = $\frac{Total\ earning}{Total\ who\ earned} = \alpha$	\$ 10 930,56 (=A)	\$ 16 591 (=a)
Eployment rate = $\frac{Total\ who\ earned}{Total\ population} = \beta$	0,717892 (=B)	0,825974 (=b)

Source: U.S. Bureau of the Census, (1984a) *Census of Population: Detailed Population Characteristics, United States Summary*, PC80-1-D1-A: Washington, D.C., table 296.

Measures	Standardization		Decomposition	
	White males (population 2)	Black males (population 1)	Difference (effects)	Percent distribution of effects
$\boldsymbol{\beta}$-standardized mean earnings	\$ 12 807,14	\$ 8 437,23	\$4 369,91 (α-effect)	74,6%
$\boldsymbol{\alpha}$-standardized mean earnings	\$ 11 365,81	\$ 9 878,55	\$1 487,72 (β-effect)	25,4%
Mean earnings (R)	\$ 13 703,73	\$ 7 846,56	\$ 5 857,17 (total effect)	100%

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

51

51

W. Das Goupta démonstration (exemple numérique 2)

Table 2.3. Crude Birth Rates as the Product of **Three Factors**: Austria and Chile, 1981

Measures	Austria, 1981 (population 1)	Chili, 1981 (population 2)
Crude birth rate = $\frac{Births \cdot 1000}{Total\ population} = R$	12,512 (= R_1)	32,845 (= R_2)
General fertility rate = $\frac{Births \cdot 1000}{Women\ aged\ 15-49} = \alpha$	51,76748 (= A)	84,90502 (= a)
Reproductive age % in women = $\frac{Women\ aged\ 15-49}{Total\ women} = \beta$	0,45919 (= B)	0,75756 (= b)
Women in population = = $\frac{Total\ women}{Total\ population} = \gamma$	0,52638 (= C)	0,51065 (= c)

Source: UN Demographic Yearbook, 1988 (table 23), 1989 (table 29).

Measures	Standardization		Decomposition	
	Chili, 1981 (population 2)	Austria, 1981 (population 1)	Difference (effects)	Percent distribution of effects
$\boldsymbol{\beta\gamma}$-standardized birth rate	26,750	16,310	10,440 (α-effect)	51,4%
$\boldsymbol{\alpha\gamma}$-standardized birth rate	26,810	16,251	10,559 (β-effect)	51,9%
$\boldsymbol{\alpha\beta}$-standardized birth rate	21,651	22,317	-0,666 (γ-effect)	-3,3%
Crude birth rate (R)	32,845	12,512	20.333 (total effect)	100%

Alexandre Avdeev, cours « Analyse démographique », IDUP, master de démographie

52

52