

Scenario comparisons: How much good can we do?

Roger B. Newson

r.newson@imperial.ac.uk

<http://www.imperial.ac.uk/nhli/r.newson/>

National Heart and Lung Institute, Imperial College London

18th UK Stata Users' Group Meeting, 13–14 September, 2012

Downloadable from the conference website at

<http://ideas.repec.org/s/boc/usug12.html>

What are scenario comparisons?

- ▶ Applied scientists, especially public health scientists, frequently want to know how much good can be caused by a proposed intervention.
- ▶ *For instance*, we might want to estimate how much we could decrease the level of a disease, in a dream scenario where the whole world stopped smoking.
- ▶ In statistics, scenarios are different versions of a dataset, with the same variables but different values and/or observations.
- ▶ We may want to compare different scenarios applied to the same population (with the same parameters).
- ▶ Alternatively, we may want to compare the same scenario between different populations (with different parameters), as when standardizing disease rates to a common distribution of gender and age.
- ▶ Scenario comparison statistics include pairwise comparisons, population attributable risks and fractions, and age-standardized heterogeneity chi-squared tests.

What are scenario comparisons?

- ▶ Applied scientists, especially public health scientists, frequently want to know how much good can be caused by a proposed intervention.
- ▶ *For instance*, we might want to estimate how much we could decrease the level of a disease, in a dream scenario where the whole world stopped smoking.
- ▶ In statistics, scenarios are different versions of a dataset, with the same variables but different values and/or observations.
- ▶ We may want to compare different scenarios applied to the same population (with the same parameters).
- ▶ Alternatively, we may want to compare the same scenario between different populations (with different parameters), as when standardizing disease rates to a common distribution of gender and age.
- ▶ Scenario comparison statistics include pairwise comparisons, population attributable risks and fractions, and age-standardized heterogeneity chi-squared tests.

What are scenario comparisons?

- ▶ Applied scientists, especially public health scientists, frequently want to know how much good can be caused by a proposed intervention.
- ▶ *For instance*, we might want to estimate how much we could decrease the level of a disease, in a dream scenario where the whole world stopped smoking.
- ▶ In statistics, scenarios are different versions of a dataset, with the same variables but different values and/or observations.
- ▶ We may want to compare different scenarios applied to the same population (with the same parameters).
- ▶ Alternatively, we may want to compare the same scenario between different populations (with different parameters), as when standardizing disease rates to a common distribution of gender and age.
- ▶ Scenario comparison statistics include pairwise comparisons, population attributable risks and fractions, and age-standardized heterogeneity chi-squared tests.

What are scenario comparisons?

- ▶ Applied scientists, especially public health scientists, frequently want to know how much good can be caused by a proposed intervention.
- ▶ *For instance*, we might want to estimate how much we could decrease the level of a disease, in a dream scenario where the whole world stopped smoking.
- ▶ In statistics, scenarios are different versions of a dataset, with the same variables but different values and/or observations.
- ▶ We may want to compare different scenarios applied to the same population (with the same parameters).
- ▶ Alternatively, we may want to compare the same scenario between different populations (with different parameters), as when standardizing disease rates to a common distribution of gender and age.
- ▶ Scenario comparison statistics include pairwise comparisons, population attributable risks and fractions, and age-standardized heterogeneity chi-squared tests.

What are scenario comparisons?

- ▶ Applied scientists, especially public health scientists, frequently want to know how much good can be caused by a proposed intervention.
- ▶ *For instance*, we might want to estimate how much we could decrease the level of a disease, in a dream scenario where the whole world stopped smoking.
- ▶ In statistics, scenarios are different versions of a dataset, with the same variables but different values and/or observations.
- ▶ We may want to compare different scenarios applied to the same population (with the same parameters).
- ▶ Alternatively, we may want to compare the same scenario between different populations (with different parameters), as when standardizing disease rates to a common distribution of gender and age.
- ▶ Scenario comparison statistics include pairwise comparisons, population attributable risks and fractions, and age-standardized heterogeneity chi-squared tests.

Existing Stata tools for scenario comparisons

- ▶ Brady (1998)[1] introduced the Stata Version 5 package `aflogit` for estimating the population attributable fraction (with confidence limits) for cohort and case–control data.
- ▶ This is still used, although it sometimes has problems with the 32–character names used in later Stata versions.
- ▶ In Stata Version 11, the `margins` command was added, allowing estimation (with confidence limits) of scenario means of a wide range of quantities.
- ▶ In Stata Version 12, the `pwcompare` command was added, together with the `pwcompare` option of `margins`, to estimate pairwise differences between scenario means.
- ▶ These new tools are very comprehensive, allowing the use of sample survey data, and the use of the Shah variance[4] to estimate confidence limits.

The `punaf` suite of packages

- ▶ The `punaf` suite of packages can be downloaded from SSC, and includes `punaf`, `punafcc`, `regpar`, `margprev` and `marglmean`.
- ▶ They use `margins` to compute confidence intervals for scenario means and proportions and/or their comparisons (differences and ratios), including population attributable (and unattributable) risks and fractions.
- ▶ These are estimated (using `nlcom`) with asymmetric confidence limits, calculated from appropriate Normalizing and variance–stabilizing transformations.
- ▶ The results may be saved as estimation results, and/or listed and/or saved in output datasets, using the SSC package `parmest`.

The `punaf` suite of packages

- ▶ The `punaf` suite of packages can be downloaded from SSC, and includes `punaf`, `punafcc`, `regpar`, `margprev` and `marglmean`.
- ▶ They use `margins` to compute confidence intervals for scenario means and proportions and/or their comparisons (differences and ratios), including population attributable (and unattributable) risks and fractions.
- ▶ These are estimated (using `nlcom`) with asymmetric confidence limits, calculated from appropriate Normalizing and variance–stabilizing transformations.
- ▶ The results may be saved as estimation results, and/or listed and/or saved in output datasets, using the SSC package `parmest`.

The `punaf` suite of packages

- ▶ The `punaf` suite of packages can be downloaded from SSC, and includes `punaf`, `punafcc`, `regpar`, `margprev` and `marglmean`.
- ▶ They use `margins` to compute confidence intervals for scenario means and proportions and/or their comparisons (differences and ratios), including population attributable (and unattributable) risks and fractions.
- ▶ These are estimated (using `nlcom`) with asymmetric confidence limits, calculated from appropriate Normalizing and variance–stabilizing transformations.
- ▶ The results may be saved as estimation results, and/or listed and/or saved in output datasets, using the SSC package `parmest`.

The `punaf` suite of packages

- ▶ The `punaf` suite of packages can be downloaded from SSC, and includes `punaf`, `punafcc`, `regpar`, `margprev` and `marglmean`.
- ▶ They use `margins` to compute confidence intervals for scenario means and proportions and/or their comparisons (differences and ratios), including population attributable (and unattributable) risks and fractions.
- ▶ These are estimated (using `nlcom`) with asymmetric confidence limits, calculated from appropriate Normalizing and variance–stabilizing transformations.
- ▶ The results may be saved as estimation results, and/or listed and/or saved in output datasets, using the SSC package `parmest`.

The `punaf` suite of packages

- ▶ The `punaf` suite of packages can be downloaded from SSC, and includes `punaf`, `punafcc`, `regpar`, `margprev` and `marglmean`.
- ▶ They use `margins` to compute confidence intervals for scenario means and proportions and/or their comparisons (differences and ratios), including population attributable (and unattributable) risks and fractions.
- ▶ These are estimated (using `nlcom`) with asymmetric confidence limits, calculated from appropriate Normalizing and variance–stabilizing transformations.
- ▶ The results may be saved as estimation results, and/or listed and/or saved in output datasets, using the SSC package `parmest`.

Packages of the `punaf` suite

These estimate and/or compare marginal means and/or prevalences for one and/or two scenarios (“Scenario 1” and “Scenario 0”), using Normalizing and variance–stabilizing transformations.

<i>Package</i>	<i>Computes confidence intervals for:</i>	<i>Transformation</i>
<code>margprev</code>	Marginal prevalences	Logit
<code>marglmean</code>	Marginal arithmetic means	Log
<code>regpar</code>	Differences between marginal prevalences (population attributable risks (PARs))	Fisher’s z
<code>punaf</code>	Ratios between marginal arithmetic means (population unattributable fractions (PUFs))	Log
<code>punafcc</code>	Arithmetic mean risk or hazard ratios (case–control or survival PUFs)	Log

A population *attributable* fraction (PAF) is estimated by end–point transformation, subtracting the corresponding PUF from 1, as recommended by Greenland and Drescher (1993)[2].

Examples in the `lbw` data

- ▶ The `lbw` dataset was discussed by Hosmer, Lemeshow and Klar (1988)[3], and is distributed on-line by Stata Press.
- ▶ It has one observation for each of a sample of 189 pregnancies, and data on the birth weight of the baby, and on a list of predictive variables.
- ▶ The most interesting of these variables is probably the mother's smoking status during pregnancy, coded as the binary variable `smoke`.
- ▶ In our examples, we will try to estimate how much good could be done by eliminating smoking (at least during pregnancy).
- ▶ This good is measured using scenario prevalences of low birth weight, which is stored in the binary variable `low` (birth weight below 2500 grams).

Examples in the `lbw` data

- ▶ The `lbw` dataset was discussed by Hosmer, Lemeshow and Klar (1988)[3], and is distributed on-line by Stata Press.
- ▶ It has one observation for each of a sample of 189 pregnancies, and data on the birth weight of the baby, and on a list of predictive variables.
- ▶ The most interesting of these variables is probably the mother's smoking status during pregnancy, coded as the binary variable `smoke`.
- ▶ In our examples, we will try to estimate how much good could be done by eliminating smoking (at least during pregnancy).
- ▶ This good is measured using scenario prevalences of low birth weight, which is stored in the binary variable `low` (birth weight below 2500 grams).

Examples in the `lbw` data

- ▶ The `lbw` dataset was discussed by Hosmer, Lemeshow and Klar (1988)[3], and is distributed on-line by Stata Press.
- ▶ It has one observation for each of a sample of 189 pregnancies, and data on the birth weight of the baby, and on a list of predictive variables.
- ▶ The most interesting of these variables is probably the mother's smoking status during pregnancy, coded as the binary variable `smoke`.
- ▶ In our examples, we will try to estimate how much good could be done by eliminating smoking (at least during pregnancy).
- ▶ This good is measured using scenario prevalences of low birth weight, which is stored in the binary variable `low` (birth weight below 2500 grams).

Examples in the `lbw` data

- ▶ The `lbw` dataset was discussed by Hosmer, Lemeshow and Klar (1988)[3], and is distributed on-line by Stata Press.
- ▶ It has one observation for each of a sample of 189 pregnancies, and data on the birth weight of the baby, and on a list of predictive variables.
- ▶ The most interesting of these variables is probably the mother's smoking status during pregnancy, coded as the binary variable `smoke`.
- ▶ In our examples, we will try to estimate how much good could be done by eliminating smoking (at least during pregnancy).
- ▶ This good is measured using scenario prevalences of low birth weight, which is stored in the binary variable `low` (birth weight below 2500 grams).

Examples in the `lbw` data

- ▶ The `lbw` dataset was discussed by Hosmer, Lemeshow and Klar (1988)[3], and is distributed on-line by Stata Press.
- ▶ It has one observation for each of a sample of 189 pregnancies, and data on the birth weight of the baby, and on a list of predictive variables.
- ▶ The most interesting of these variables is probably the mother's smoking status during pregnancy, coded as the binary variable `smoke`.
- ▶ In our examples, we will try to estimate how much good could be done by eliminating smoking (at least during pregnancy).
- ▶ This good is measured using scenario prevalences of low birth weight, which is stored in the binary variable `low` (birth weight below 2500 grams).

Logistic regression in the lbw data

After loading the lbw data, we fit a logistic regression model of low with respect to smoke and the confounder race (white, black or other):

```
. logit low i.race i.smoke, or vce(robust);
```

```
Iteration 0:  log pseudolikelihood =  -117.336
Iteration 1:  log pseudolikelihood = -110.10441
Iteration 2:  log pseudolikelihood = -109.98749
Iteration 3:  log pseudolikelihood = -109.98736
Iteration 4:  log pseudolikelihood = -109.98736
```

```
Logistic regression                               Number of obs   =       189
                                                    Wald chi2(3)    =       14.30
                                                    Prob > chi2     =       0.0025
Log pseudolikelihood = -109.98736                Pseudo R2      =       0.0626
```

	low	Odds Ratio	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
race							
2		2.956742	1.420439	2.26	0.024	1.153162	7.581175
3		3.030001	1.187272	2.83	0.005	1.405753	6.530954
1.smoke		3.052631	1.10296	3.09	0.002	1.503568	6.197631
_cons		.1587319	.0515235	-5.67	0.000	.0840173	.2998882

We see that maternal smoking trebles the odds of low birth weight.

But how much good can we do?

- ▶ Not many people really understand odds ratios.
- ▶ An audience of non-mathematicians might prefer to know what difference it would make, if all pregnant mothers stopped smoking.
- ▶ The `regpar` package can answer this question, by comparing prevalences of low birth rate under 2 scenarios.
- ▶ “Scenario 0” is the world we live in, where some mothers smoke.
- ▶ “Scenario 1” is a fantasy world, where no mothers smoke, but the race distribution stays the same.
- ▶ The difference between these prevalences is the **population attributable risk (PAR)**.

But how much good can we do?

- ▶ Not many people really understand odds ratios.
- ▶ An audience of non-mathematicians might prefer to know what difference it would make, if all pregnant mothers stopped smoking.
- ▶ The `regpar` package can answer this question, by comparing prevalences of low birth rate under 2 scenarios.
- ▶ “Scenario 0” is the world we live in, where some mothers smoke.
- ▶ “Scenario 1” is a fantasy world, where no mothers smoke, but the race distribution stays the same.
- ▶ The difference between these prevalences is the **population attributable risk (PAR)**.

Scenario prevalences and the PAR using regpar

We execute `regpar`, as follows:

```
. regpar, at(smoke=0);
Scenario 0: (asobserved) _all
Scenario 1: smoke=0
Symmetric confidence intervals for the logit proportions
under Scenario 0 and Scenario 1
and for the z-transformed population attributable risk (PAR)
Total number of observations used: 189
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Scenario_0	-.789997	.1519305	-5.20	0.000	-1.087775	-.4922187
Scenario_1	-1.215955	.2051031	-5.93	0.000	-1.61795	-.8139606
PAR	.0837153	.0266196	3.14	0.002	.0315419	.1358887

```
Asymmetric 95% CIs for the untransformed proportions
under Scenario 0 and Scenario 1
and for the untransformed population attributable risk (PAR)
```

	Estimate	Minimum	Maximum
Scenario_0	.31216931	.25203743	.37937104
Scenario_1	.22864901	.16548776	.30704715
PAR	.08352031	.03153146	.13505843

We see that 3.2 to 13.5 percent of all babies might be saved from low birth weight, if no mothers smoked.

Advice for smoking mothers

- ▶ Our *real* aim is to communicate our message to an audience of smoking mothers, and not just to an audience of target-minded public health professionals.
- ▶ This non-professional audience might want to know what good *they* could do for *their* babies.
- ▶ The `regpar` package can answer this question, too, by comparing prevalences of low birth weight under the 2 scenarios *in the subpopulation of smoking mothers*.
- ▶ This is done using the `subpop()` option of `regpar`.

Advice for smoking mothers

- ▶ Our *real* aim is to communicate our message to an audience of smoking mothers, and not just to an audience of target-minded public health professionals.
- ▶ This non-professional audience might want to know what good *they* could do for *their* babies.
- ▶ The `regpar` package can answer this question, too, by comparing prevalences of low birth weight under the 2 scenarios *in the subpopulation of smoking mothers*.
- ▶ This is done using the `subpop()` option of `regpar`.

Advice for smoking mothers

- ▶ Our *real* aim is to communicate our message to an audience of smoking mothers, and not just to an audience of target-minded public health professionals.
- ▶ This non-professional audience might want to know what good *they* could do for *their* babies.
- ▶ The `regpar` package can answer this question, too, by comparing prevalences of low birth weight under the 2 scenarios *in the subpopulation of smoking mothers*.
- ▶ This is done using the `subpop()` option of `regpar`.

Advice for smoking mothers

- ▶ Our *real* aim is to communicate our message to an audience of smoking mothers, and not just to an audience of target-minded public health professionals.
- ▶ This non-professional audience might want to know what good *they* could do for *their* babies.
- ▶ The `regpar` package can answer this question, too, by comparing prevalences of low birth weight under the 2 scenarios *in the subpopulation of smoking mothers*.
- ▶ This is done using the `subpop()` option of `regpar`.

Scenario prevalences and the exposed subpopulation attributable risk for smoking mothers

We execute `regpar` with the `subpop()` option:

```
. regpar, at(smoke=0) subpop(if smoke==1);
Scenario 0: (asobserved) _all
Scenario 1: smoke=0
Symmetric confidence intervals for the logit proportions
under Scenario 0 and Scenario 1
and for the z-transformed population attributable risk (PAR)
Total number of observations used: 189
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Scenario_0	-.3829923	.2373852	-1.61	0.107	-.8482587	.0822742
Scenario_1	-1.436486	.2279922	-6.30	0.000	-1.883343	-.9896299
PAR	.2166422	.0707321	3.06	0.002	.0780098	.3552746

```
Asymmetric 95% CIs for the untransformed proportions
under Scenario 0 and Scenario 1
and for the untransformed population attributable risk (PAR)
```

	Estimate	Minimum	Maximum
Scenario_0	.40540541	.29979827	.52055695
Scenario_1	.19209003	.13200536	.27098519
PAR	.21331537	.07785194	.34104503

We see that 7.8 to 34.1 percent of babies *of smoking mothers* might be saved, if none of their mothers smoked.

Attributable disease burden as a fraction of the total disease burden

- ▶ Returning to our previous audience of target-minded professionals, we might be asked what percent of the “burden” of low birthweight might be removed by eliminating smoking.
- ▶ The package to answer this is `puna`, which calculates population unattributable and attributable *fractions*.
- ▶ Prevalences and rates are arithmetic means of non-negative variables.
- ▶ `puna` estimates 2 scenario means of the same non-negative variable, and their ratio, the population *unattributable* fraction.
- ▶ `puna` then subtracts this ratio from 1 to get the **population attributable fraction**.

Attributable disease burden as a fraction of the total disease burden

- ▶ Returning to our previous audience of target-minded professionals, we might be asked what percent of the “burden” of low birthweight might be removed by eliminating smoking.
- ▶ The package to answer this is `puna`, which calculates population unattributable and attributable *fractions*.
- ▶ Prevalences and rates are arithmetic means of non-negative variables.
- ▶ `puna` estimates 2 scenario means of the same non-negative variable, and their ratio, the population *unattributable* fraction.
- ▶ `puna` then subtracts this ratio from 1 to get the **population attributable fraction**.

Attributable disease burden as a fraction of the total disease burden

- ▶ Returning to our previous audience of target-minded professionals, we might be asked what percent of the “burden” of low birthweight might be removed by eliminating smoking.
- ▶ The package to answer this is `puna`, which calculates population unattributable and attributable *fractions*.
- ▶ Prevalences and rates are arithmetic means of non-negative variables.
- ▶ `puna` estimates 2 scenario means of the same non-negative variable, and their ratio, the population *unattributable* fraction.
- ▶ `puna` then subtracts this ratio from 1 to get the **population attributable fraction**.

Attributable disease burden as a fraction of the total disease burden

- ▶ Returning to our previous audience of target-minded professionals, we might be asked what percent of the “burden” of low birthweight might be removed by eliminating smoking.
- ▶ The package to answer this is `puna`, which calculates population unattributable and attributable *fractions*.
- ▶ Prevalences and rates are arithmetic means of non-negative variables.
- ▶ `puna` estimates 2 scenario means of the same non-negative variable, and their ratio, the population *unattributable* fraction.
- ▶ `puna` then subtracts this ratio from 1 to get the **population attributable fraction**.

Attributable disease burden as a fraction of the total disease burden

- ▶ Returning to our previous audience of target-minded professionals, we might be asked what percent of the “burden” of low birthweight might be removed by eliminating smoking.
- ▶ The package to answer this is `puna`, which calculates population unattributable and attributable *fractions*.
- ▶ Prevalences and rates are arithmetic means of non-negative variables.
- ▶ `puna` estimates 2 scenario means of the same non-negative variable, and their ratio, the population *unattributable* fraction.
- ▶ `puna` then subtracts this ratio from 1 to get the **population attributable fraction**.

Scenario prevalences and the population unattributable and attributable fractions

We execute `punaf` as follows:

```
. punaf, at(smoke=0) eform;
Scenario 0: (asobserved) _all
Scenario 1: smoke=0
Confidence intervals for the means under Scenario 0 and Scenario 1
and for the population unattributable fraction (PUF)
Total number of observations used: 189
-----
```

	Mean/Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
Scenario_0	.3121693	.0326225	-11.14	0.000	.2543534	.3831271
Scenario_1	.228649	.0361738	-9.33	0.000	.1676887	.3117704
PUF	.7324519	.0818807	-2.79	0.005	.5883333	.911874

```
-----
```

```
95% CI for the population attributable fraction (PAF)
      Estimate      Minimum      Maximum
PAF      .2675481      .08812601      .41166675
```

The scenario prevalences are estimated, with their ratio, the PUF. This is subtracted from 1 to estimate the PAF, which is 8.8 to 41.2 percent of the Scenario 0 prevalence.

Attributable burden as a fraction of the total burden from smoking mothers

- ▶ For completeness, we can estimate the *exposed subpopulation* attributable fraction.
- ▶ This is done using `punaaf`, with a `subpop()` option.

Attributable burden as a fraction of the total burden from smoking mothers

- ▶ For completeness, we can estimate the *exposed subpopulation* attributable fraction.
- ▶ This is done using `puna_f`, with a `subpop()` option.

Attributable burden as a fraction of the total burden from smoking mothers

- ▶ For completeness, we can estimate the *exposed subpopulation* attributable fraction.
- ▶ This is done using `punaaf`, with a `subpop()` option.

Scenario prevalences and the exposed subpopulation unattributable and attributable fractions

We execute `punaf` as follows:

```
. punaf, at(smoke=0) subpop(if smoke==1) eform;
```

```
Scenario 0: (asobserved) _all
```

```
Scenario 1: smoke=0
```

```
Confidence intervals for the means under Scenario 0 and Scenario 1  
and for the population unattributable fraction (PUF)
```

```
Total number of observations used: 189
```

	Mean/Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
Scenario_0	.4054054	.0572221	-6.40	0.000	.3074285	.5346073
Scenario_1	.19209	.0353824	-8.96	0.000	.1338801	.2756092
PUF	.4738221	.1103706	-3.21	0.001	.3001505	.7479826

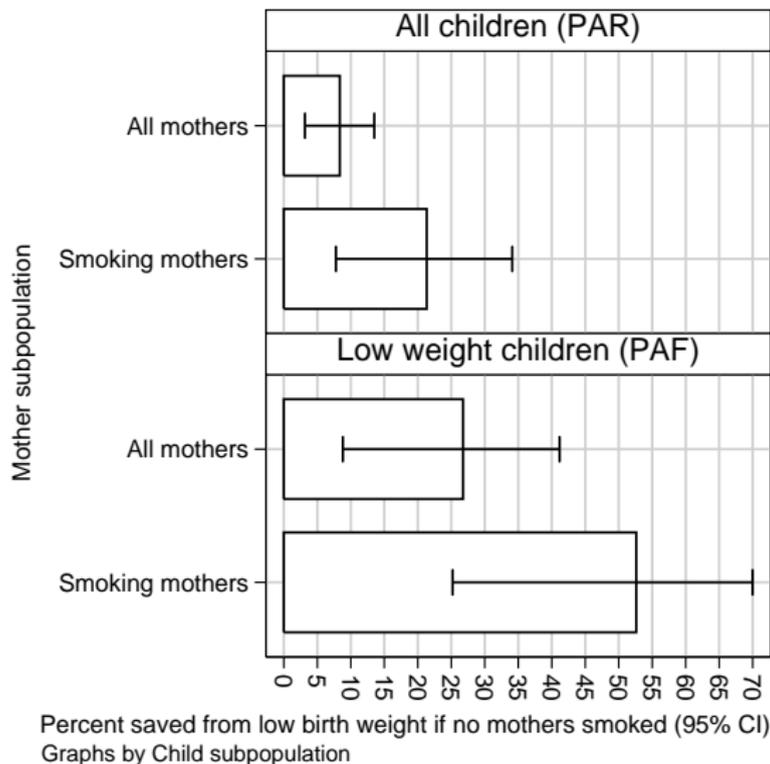
```
95% CI for the population attributable fraction (PAF)
```

	Estimate	Minimum	Maximum
PAF	.52617792	.25201736	.6998495

This time, we see that 25.2 to 70.0 percent of the low-weight births *from smoking mothers* can be attributed to smoking.

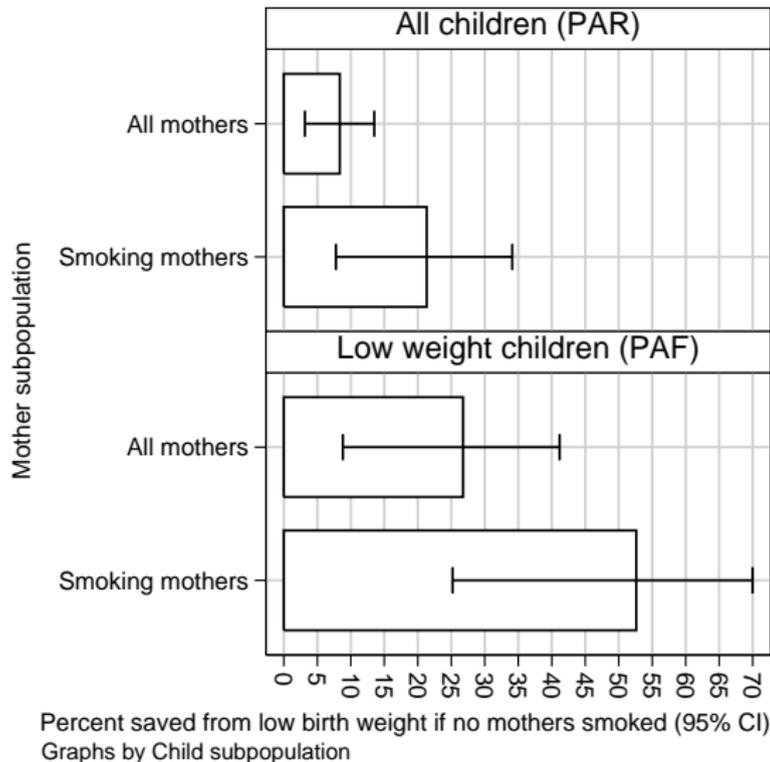
Summary: Attributable risks and fractions in the 1bw data

- ▶ The upper plot gives PARs, measuring prevention in all children.
- ▶ The lower plot gives PAFs, measuring prevention in low birth weight children.
- ▶ In both cases, proportions prevented are greater for children with smoking mothers.
- ▶ These conclusions assume that the association is causal.



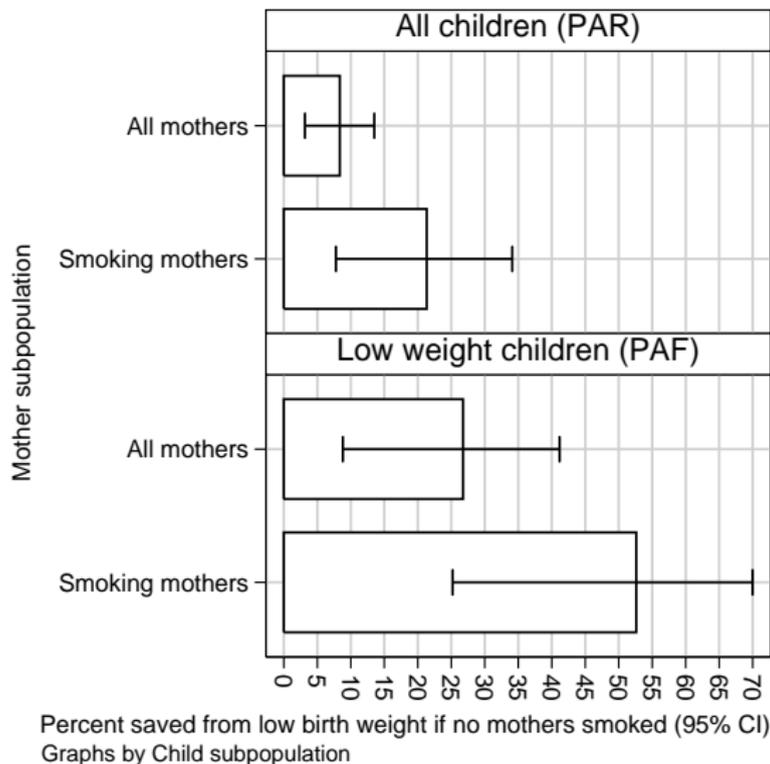
Summary: Attributable risks and fractions in the 1bw data

- ▶ The upper plot gives PARs, measuring prevention in all children.
- ▶ The lower plot gives PAFs, measuring prevention in low birth weight children.
- ▶ In both cases, proportions prevented are greater for children with smoking mothers.
- ▶ These conclusions assume that the association is causal.



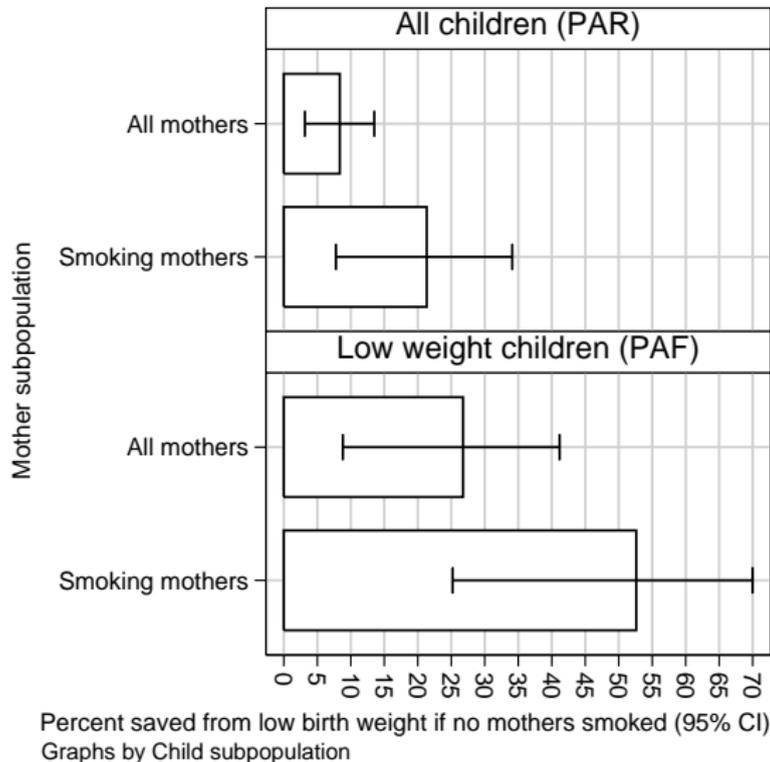
Summary: Attributable risks and fractions in the 1bw data

- ▶ The upper plot gives PARs, measuring prevention in all children.
- ▶ The lower plot gives PAFs, measuring prevention in low birth weight children.
- ▶ In both cases, proportions prevented are greater for children with smoking mothers.
- ▶ These conclusions assume that the association is causal.



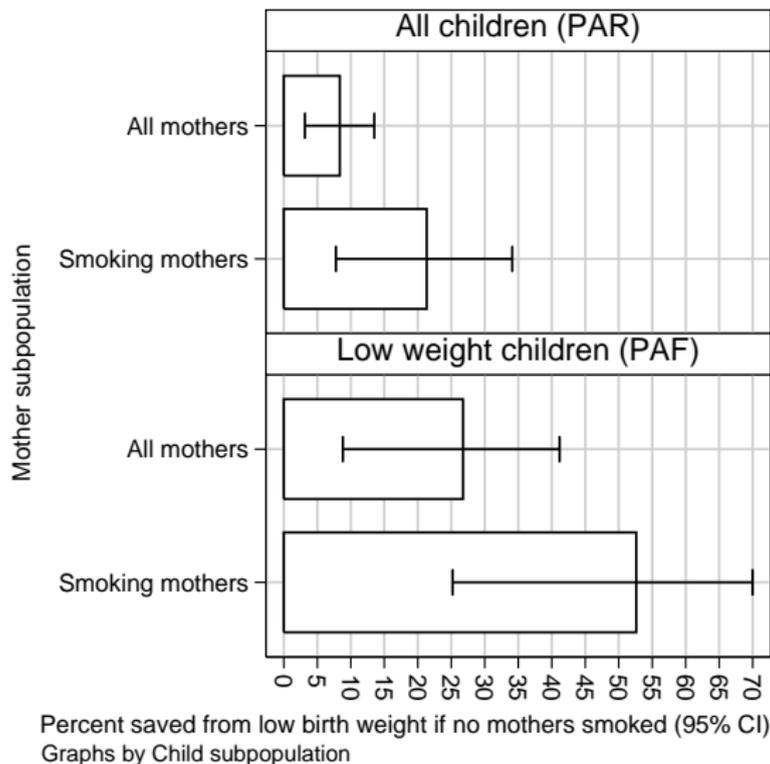
Summary: Attributable risks and fractions in the 1bw data

- ▶ The upper plot gives PARs, measuring prevention in all children.
- ▶ The lower plot gives PAFs, measuring prevention in low birth weight children.
- ▶ In both cases, proportions prevented are greater for children with smoking mothers.
- ▶ These conclusions assume that the association is causal.



Summary: Attributable risks and fractions in the 1bw data

- ▶ The upper plot gives PARs, measuring prevention in all children.
- ▶ The lower plot gives PAFs, measuring prevention in low birth weight children.
- ▶ In both cases, proportions prevented are greater for children with smoking mothers.
- ▶ These conclusions assume that the association is causal.



Standardization as out-of-sample prediction

- ▶ The `punaf` suite may also be used to compare the *same* scenario between *different* models, as well as *vice versa*.
- ▶ For example, we might fit multiple logit models to multiple independent subpopulation datasets, and then estimate the marginal prevalence that each model would imply in a standard-population dataset, with a standard distribution of gender and age.
- ▶ This practice is an example of **out-of-sample prediction**.
- ▶ To do this, packages of the `punaf` suite all have an option `noesample`, functioning as the option of the same name for `margins`.
- ▶ We will illustrate this with an example using `margprev`.

Standardization as out-of-sample prediction

- ▶ The `puna` suite may also be used to compare the *same* scenario between *different* models, as well as *vice versa*.
- ▶ For example, we might fit multiple logit models to multiple independent subpopulation datasets, and then estimate the marginal prevalence that each model would imply in a standard-population dataset, with a standard distribution of gender and age.
- ▶ This practice is an example of **out-of-sample prediction**.
- ▶ To do this, packages of the `puna` suite all have an option `noesample`, functioning as the option of the same name for `margins`.
- ▶ We will illustrate this with an example using `margprev`.

Standardization as out-of-sample prediction

- ▶ The `puna` suite may also be used to compare the *same* scenario between *different* models, as well as *vice versa*.
- ▶ For example, we might fit multiple logit models to multiple independent subpopulation datasets, and then estimate the marginal prevalence that each model would imply in a standard-population dataset, with a standard distribution of gender and age.
- ▶ This practice is an example of **out-of-sample prediction**.
- ▶ To do this, packages of the `puna` suite all have an option `noesample`, functioning as the option of the same name for `margins`.
- ▶ We will illustrate this with an example using `margprev`.

Standardization as out-of-sample prediction

- ▶ The `puna` suite may also be used to compare the *same* scenario between *different* models, as well as *vice versa*.
- ▶ For example, we might fit multiple logit models to multiple independent subpopulation datasets, and then estimate the marginal prevalence that each model would imply in a standard-population dataset, with a standard distribution of gender and age.
- ▶ This practice is an example of **out-of-sample prediction**.
- ▶ To do this, packages of the `puna` suite all have an option `noesample`, functioning as the option of the same name for `margins`.
- ▶ We will illustrate this with an example using `margprev`.

Standardization as out-of-sample prediction

- ▶ The `punaf` suite may also be used to compare the *same* scenario between *different* models, as well as *vice versa*.
- ▶ For example, we might fit multiple logit models to multiple independent subpopulation datasets, and then estimate the marginal prevalence that each model would imply in a standard-population dataset, with a standard distribution of gender and age.
- ▶ This practice is an example of **out-of-sample prediction**.
- ▶ To do this, packages of the `punaf` suite all have an option `noesample`, functioning as the option of the same name for `margins`.
- ▶ We will illustrate this with an example using `margprev`.

Skin prick allergy prevalences in the GA²LEN Follow-up Survey

- ▶ The GA²LEN Follow-up Survey is part of a multi-centre European study on asthma and allergy.
- ▶ Sensitivity to a number of allergens was measured, using skin prick tests, in 13 regional subpopulations distributed over Europe.
- ▶ For each allergen in each subpopulation, we fitted a logit model of sensitivity to that allergen in that subpopulation, with respect to age and gender.
- ▶ The parameters estimated for each such model were then used to estimate the marginal prevalence expected, if this model applied to a European Standard Population with a standard gender and age distribution.
- ▶ This was done using `margprev`, with the `noesample` option, in a dataset representing the European Standard Population.
- ▶ These standardized marginal prevalences were then compared between regional subpopulations, using a chi-squared heterogeneity test, to detect between-region heterogeneity.

Skin prick allergy prevalences in the GA²LEN Follow-up Survey

- ▶ The GA²LEN Follow-up Survey is part of a multi-centre European study on asthma and allergy.
- ▶ Sensitivity to a number of allergens was measured, using skin prick tests, in 13 regional subpopulations distributed over Europe.
- ▶ For each allergen in each subpopulation, we fitted a logit model of sensitivity to that allergen in that subpopulation, with respect to age and gender.
- ▶ The parameters estimated for each such model were then used to estimate the marginal prevalence expected, if this model applied to a European Standard Population with a standard gender and age distribution.
- ▶ This was done using `margprev`, with the `noesample` option, in a dataset representing the European Standard Population.
- ▶ These standardized marginal prevalences were then compared between regional subpopulations, using a chi-squared heterogeneity test, to detect between-region heterogeneity.

Skin prick allergy prevalences in the GA²LEN Follow-up Survey

- ▶ The GA²LEN Follow-up Survey is part of a multi-centre European study on asthma and allergy.
- ▶ Sensitivity to a number of allergens was measured, using skin prick tests, in 13 regional subpopulations distributed over Europe.
- ▶ For each allergen in each subpopulation, we fitted a logit model of sensitivity to that allergen in that subpopulation, with respect to age and gender.
- ▶ The parameters estimated for each such model were then used to estimate the marginal prevalence expected, if this model applied to a European Standard Population with a standard gender and age distribution.
- ▶ This was done using `margprev`, with the `noesample` option, in a dataset representing the European Standard Population.
- ▶ These standardized marginal prevalences were then compared between regional subpopulations, using a chi-squared heterogeneity test, to detect between-region heterogeneity.

Skin prick allergy prevalences in the GA²LEN Follow-up Survey

- ▶ The GA²LEN Follow-up Survey is part of a multi-centre European study on asthma and allergy.
- ▶ Sensitivity to a number of allergens was measured, using skin prick tests, in 13 regional subpopulations distributed over Europe.
- ▶ For each allergen in each subpopulation, we fitted a logit model of sensitivity to that allergen in that subpopulation, with respect to age and gender.
- ▶ The parameters estimated for each such model were then used to estimate the marginal prevalence expected, if this model applied to a European Standard Population with a standard gender and age distribution.
- ▶ This was done using `margprev`, with the `noesample` option, in a dataset representing the European Standard Population.
- ▶ These standardized marginal prevalences were then compared between regional subpopulations, using a chi-squared heterogeneity test, to detect between-region heterogeneity.

Skin prick allergy prevalences in the GA²LEN Follow-up Survey

- ▶ The GA²LEN Follow-up Survey is part of a multi-centre European study on asthma and allergy.
- ▶ Sensitivity to a number of allergens was measured, using skin prick tests, in 13 regional subpopulations distributed over Europe.
- ▶ For each allergen in each subpopulation, we fitted a logit model of sensitivity to that allergen in that subpopulation, with respect to age and gender.
- ▶ The parameters estimated for each such model were then used to estimate the marginal prevalence expected, if this model applied to a European Standard Population with a standard gender and age distribution.
- ▶ This was done using `margprev`, with the `noesample` option, in a dataset representing the European Standard Population.
- ▶ These standardized marginal prevalences were then compared between regional subpopulations, using a chi-squared heterogeneity test, to detect between-region heterogeneity.

Skin prick allergy prevalences in the GA²LEN Follow-up Survey

- ▶ The GA²LEN Follow-up Survey is part of a multi-centre European study on asthma and allergy.
- ▶ Sensitivity to a number of allergens was measured, using skin prick tests, in 13 regional subpopulations distributed over Europe.
- ▶ For each allergen in each subpopulation, we fitted a logit model of sensitivity to that allergen in that subpopulation, with respect to age and gender.
- ▶ The parameters estimated for each such model were then used to estimate the marginal prevalence expected, if this model applied to a European Standard Population with a standard gender and age distribution.
- ▶ This was done using `margprev`, with the `noesample` option, in a dataset representing the European Standard Population.
- ▶ These standardized marginal prevalences were then compared between regional subpopulations, using a chi-squared heterogeneity test, to detect between-region heterogeneity.

Skin prick allergy prevalences in the GA²LEN Follow-up Survey

- ▶ The GA²LEN Follow-up Survey is part of a multi-centre European study on asthma and allergy.
- ▶ Sensitivity to a number of allergens was measured, using skin prick tests, in 13 regional subpopulations distributed over Europe.
- ▶ For each allergen in each subpopulation, we fitted a logit model of sensitivity to that allergen in that subpopulation, with respect to age and gender.
- ▶ The parameters estimated for each such model were then used to estimate the marginal prevalence expected, if this model applied to a European Standard Population with a standard gender and age distribution.
- ▶ This was done using `margprev`, with the `noesample` option, in a dataset representing the European Standard Population.
- ▶ These standardized marginal prevalences were then compared between regional subpopulations, using a chi-squared heterogeneity test, to detect between-region heterogeneity.

Logit regression of cat allergy with respect to gender and age in the UK

In the UK subpopulation, we fitted a logit model of cat allergy with respect to gender and age, using sampling–probability weights:

```
. logit spt_cat male fquesagec [pwei=sampwt5], or;
```

```
Iteration 0:  log pseudolikelihood = -1030.8768
Iteration 1:  log pseudolikelihood = -977.80033
Iteration 2:  log pseudolikelihood = -973.41056
Iteration 3:  log pseudolikelihood = -973.39866
Iteration 4:  log pseudolikelihood = -973.39866
```

```
Logistic regression                               Number of obs   =          159
                                                    Wald chi2(2)    =           4.04
                                                    Prob > chi2     =          0.1328
Log pseudolikelihood = -973.39866                Pseudo R2      =          0.0558
```

```
-----+-----
            |                   Robust
spt_cat    | Odds Ratio   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      male |  2.527963   1.535882     1.53  0.127    .7684525   8.316188
fquesagec |  .6700974   .2209261    -1.21  0.225    .3511585   1.278712
   _cons  |  .0794547   .0300632    -6.69  0.000    .0378487   .1667967
-----+-----
```

The parameter `_cons` is the odds of cat allergy in 48–year–old women. The others are male–gender and per–decade odds ratios. *However*, not many people understand those parameters. *So...*

Dataset representing a European Standard Population

... we now load a dataset with 1 observation per gender per age group, and data on the gender, the minimum and maximum ages in the age group, the mean age (uncentered in years and centered at 48 in decades), and the number in that gender and age group in a European Standard Population:

```
. use stanpop, clear;  
. list male agemin agemax agemean fquesagec stanpop, abbr(32) sepby(male);
```

	male	agemin	agemax	agemean	fquesagec	stanpop
1.	0	20	24	22	-2.6	7000
2.	0	25	29	27	-2.1	7000
3.	0	30	34	32	-1.6	7000
4.	0	35	39	37	-1.1	7000
5.	0	40	44	42	-.6	7000
6.	0	45	49	47	-.1	7000
7.	0	50	54	52	.4	7000
8.	0	55	59	57	.9	6000

(Here, we have only showed the younger female age groups.)

Cat allergy prevalence standardized to the European Standard Population

Having loaded the dataset, we now use `margprev`, with frequency weights and the `noesample` option, to estimate the marginal odds and prevalence of cat allergy in the European Standard Population, using parameters from the model we fitted earlier for the UK:

```
. margprev [fwei=stanpop], eform noesample;
Scenario 1: (asobserved) _all
Confidence interval for the marginal odds
under Scenario 1
Total number of observations used: 134000
```

	Odds	Std. Err.	z	P> z	[95% Conf. Interval]
Scenario_1	.1782219	.07486	-4.11	0.000	.0782391 .4059742

```
Asymmetric 95% CI for the untransformed marginal prevalence
under Scenario 1
```

	Estimate	Minimum	Maximum
Scenario_1	.15126346	.07256191	.2887494

The standardized odds for the UK subpopulation, and for the 12 other regional subpopulations, were input into the SSC package `parmhet` to test for heterogeneity of cat allergy prevalence between European regions. This heterogeneity was detectable ($I^2=46.4\%$, $P=.033$).

