Nonparametric Estimation of the Potential Impact Fraction and the Population Attributable Fraction

Colleen Chan

Yale University, Dept. of Statistics and Data Science

Collaborators: Rodrigo Zepeda-Tello[1], Dalia Camacho García Formentí[1], Frederick Cudhea[2], Rafael Meza[3], Eliane Rodrigues[4], Tonatiuh Barrientos Gutierrez[1], Donna Spiegelman[5], Xin Zhou[5]

[1]National Institute of Public Health of Mexico, [2]Tufts University Friedman School of Nutrition Science and Policy, [3]University of Michigan School of Public Health, [4]Universidad Nacional Autonoma de Mexico, [5]Yale School of Public Health

Paper link: https://arxiv.org/abs/2207.03597

IBC 2022 Young Statistician's Showcase

12 July, 2022

SSB's and Type 2 Diabetes

Sugar-sweetened beverages (SSB's)

- Drinks with added sugar
- The largest source of added sugar in our diets today. SSB intake has risen most dramatically in LMIC's[1]
- SSB consumption linked to increased risk of T2D, obesity, heart disease

[1] Malik et al., *Nature Reviews Endocrinology*, 2022.

SSB's and Type 2 Diabetes

Sugar-sweetened beverages (SSB's)

- Drinks with added sugar
- The largest source of added sugar in our diets today. SSB intake has risen most dramatically in LMIC's[1]
- SSB consumption linked to increased risk of T2D, obesity, heart disease

Q: What fraction of type 2 diabetes cases can be attributed to SSB consumption? What if SSB consumption were entirely eliminated? What if it were halved?

[1]Malik et al., *Nature Reviews Endocrinology*, 2022.

The PIF and PAF

- The potential impact fraction (PIF), or the attributable fraction, is the proportion of incidents attributable to a given risk factor
- It requires a relative risk (RR) function that depends on exposure levels $\boldsymbol{X}$ and regression coefficients β
  - Most common form $RR(\boldsymbol{X};\beta) = \exp(\boldsymbol{X}\beta)$

The PIF and PAF

- The potential impact fraction (PIF), or the attributable fraction, is the proportion of incidents attributable to a given risk factor
- It requires a relative risk (RR) function that depends on exposure levels $\boldsymbol{X}$ and regression coefficients β
  - Most common form $RR(\boldsymbol{X};\beta) = \exp(\boldsymbol{X}\beta)$

Definition

The **potential impact fraction (PIF)** is defined as

$$\text{PIF} = \frac{\mathbb{E}_{\boldsymbol{X}}^{\text{obs}}\Big[RR\big(\boldsymbol{X};\beta\big)\Big] - \mathbb{E}_{\boldsymbol{X}}^{\text{cft}}\Big[RR\big(\boldsymbol{X};\beta\big)\Big]}{\mathbb{E}_{\boldsymbol{X}}^{\text{obs}}\Big[RR\big(\boldsymbol{X};\beta\big)\Big]}, \tag{1}$$

where $\mathbb{E}_{\boldsymbol{X}}^{\text{obs}}[RR(\boldsymbol{X};\beta)]$ represents the expected value of the relative risk under the observed exposure distribution and $\mathbb{E}_{\boldsymbol{X}}^{\text{cft}}[RR(\boldsymbol{X};\beta)]$ is the expected value of the relative risk under a counterfactual distribution of the exposure.

The PIF and PAF

- The population attributable fraction (PAF), or the attributable fraction for the population, is a specific case of the PIF when the counterfactual exposure is 0 ($\mathbb{E}^{\text{cft}}_{\boldsymbol{X}}\left[RR(\boldsymbol{X};\beta)\right] = 1$)

Definition

The **population attributable fraction (PAF)** is defined as

$$\text{PAF} = 1 - \frac{1}{\mathbb{E}^{\text{obs}}_{\boldsymbol{X}}\left[RR\left(\boldsymbol{X};\beta\right)\right]}, \tag{2}$$

where $\mathbb{E}^{\text{obs}}_{\boldsymbol{X}}\left[RR(\boldsymbol{X};\beta)\right]$ represents the expected value of the relative risk under the observed exposure distribution in a given population.

Standard Approach[1]

1. Assume a parametric distribution for continuous exposure $\boldsymbol{X}$ (e.g. Log normal, Weibull, Gamma)
2. Fit the parameters using method of moments estimation, matching the mean and variance of the observed exposure data
3. Estimate the PIF from Eq. 1 or the PAF from Eq. 2 using analytic or numerical integration

[1] GBD 2013 Risk Factors et al., 2015, Gortmaker et al., 2016, Veerman et al., 2016

Standard Approach[1]

1. Assume a parametric distribution for continuous exposure $\boldsymbol{X}$ (e.g. Log normal, Weibull, Gamma)
2. Fit the parameters using method of moments estimation, matching the mean and variance of the observed exposure data
3. Estimate the PIF from Eq. 1 or the PAF from Eq. 2 using analytic or numerical integration

Issues with the standard approach:

1. PIF is undefined for heavy-tailed exposure distributions
2. PIF can be heavily biased if exposure distribution is misspecified

[1] GBD 2013 Risk Factors et al., 2015, Gortmaker et al., 2016, Veerman et al., 2016

Undefined PIF's?

$$\text{PAF} = 1 - \frac{1}{\mathbb{E}_{\boldsymbol{X}}^{\text{obs}}\left[RR\left(\boldsymbol{X};\boldsymbol{\beta}\right)\right]}$$

- The problem lies on the combination of a heavy-tailed distribution with an exponential relative risk
- A random variable X is said to have a heavy tail if the tail probabilities $P(X > t)$ decay more slowly than tails of any exponential distribution

$$\lim_{x\to\infty} e^{cx} P(X > x) = \infty \text{ for all positive } c$$

- Distributions with heavy tails: Log normal, Pareto, Cauchy, Weibull with shape parameter less than 1

Standard Approach

Table: Relative bias percentage of PAF under different distributional assumptions for the standard method.

		Distribution assumed			
True distribution	True PAF	Gamma	Log normal	Normal	Weibull
Gamma(1.15, 1.29)	0.3455				
Normal(1.48, 1.38)	0.3795				
Weibull(1.08, 1.53)	0.3447				

Standard Approach

Table: Relative bias percentage of PAF under different distributional assumptions for the standard method.

		Distribution assumed			
True distribution	True PAF	Gamma	Log normal	Normal	Weibull
Gamma(1.15, 1.29)	0.3455	0			
Normal(1.48, 1.38)	0.3795			0	
Weibull(1.08, 1.53)	0.3447				0

Standard Approach

Table: Relative bias percentage of PAF under different distributional assumptions for the standard method.

		Distribution assumed			
True distribution	True PAF	Gamma	Log normal	Normal	Weibull
Gamma(1.15, 1.29)	0.3455	0	189.4	-19.6	-0.2
Normal(1.48, 1.38)	0.3795	-9.2	163.5	0	-9.3
Weibull(1.08, 1.53)	0.3447	0.2	190.1	-19.2	0

(Kehoe) Mixture Approach

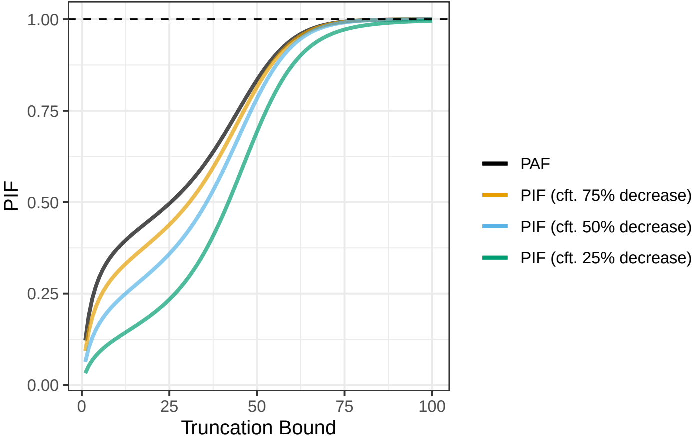
To avoid undefined PIF values, Kehoe et al. (2012) proposes:

- Truncate the assumed exposure distribution by an upper bound M
- Fit the exposure data using maximum likelihood estimation. Separate out 0 and positive values of the exposure

$$\text{PAF} = 1 - \frac{1}{p_0 RR_0 + \int_0^M RR(\boldsymbol{X};\beta) f(\boldsymbol{X}) \mathrm{d}\boldsymbol{X}}$$

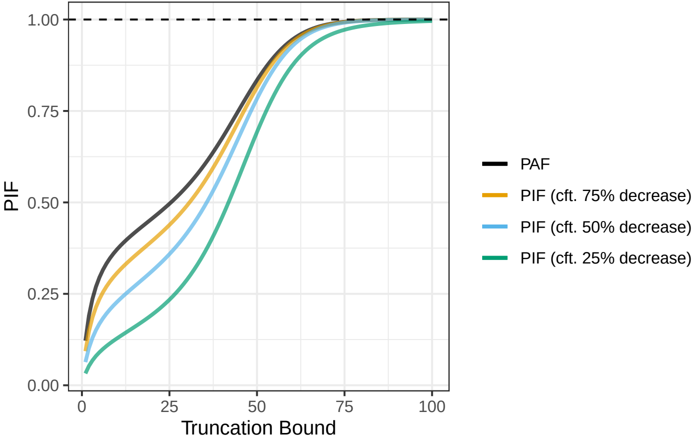
Mixture Approach

PIF value now depends on truncation bound!



Mixture Approach

PIF value now depends on truncation bound!



We propose two nonparametric methods: empirical method and approximate method

Methods - Empirical Method

Let $\hat{\mu}_n^{\mathrm{obs}}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n} RR\big(\boldsymbol{X}_i; \boldsymbol{\beta}\big)$ and $\hat{\mu}_n^{\mathrm{cft}}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n} RR\big(g(\boldsymbol{X}_i); \boldsymbol{\beta}\big)$.
We define the empirical estimators of the PAF and PIF as:

$$\widehat{\mathrm{PAF}} := 1 - \frac{1}{\hat{\mu}_n^{\mathrm{obs}}(\hat{\boldsymbol{\beta}})}, \quad \text{and} \quad \widehat{\mathrm{PIF}} := 1 - \frac{\hat{\mu}_n^{\mathrm{cft}}(\hat{\boldsymbol{\beta}})}{\hat{\mu}_n^{\mathrm{obs}}(\hat{\boldsymbol{\beta}})}.$$

Methods - Empirical Method

Let $\hat{\mu}_n^{\mathrm{obs}}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n} RR\big(\boldsymbol{X}_i;\boldsymbol{\beta}\big)$ and $\hat{\mu}_n^{\mathrm{cft}}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n} RR\big(g(\boldsymbol{X}_i);\boldsymbol{\beta}\big)$.
We define the empirical estimators of the PAF and PIF as:

$$\widehat{\mathsf{PAF}} := 1 - \frac{1}{\hat{\mu}_n^{\mathrm{obs}}(\hat{\boldsymbol{\beta}})}, \quad \text{and} \quad \widehat{\mathsf{PIF}} := 1 - \frac{\hat{\mu}_n^{\mathrm{cft}}(\hat{\boldsymbol{\beta}})}{\hat{\mu}_n^{\mathrm{obs}}(\hat{\boldsymbol{\beta}})}.$$

Theorem

Suppose that $\hat{\boldsymbol{\beta}}$ is a consistent and asymptotically normal estimator from an independent study. That is, $\sqrt{m}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is asymptotically mean-zero multivariate normal with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$, where m is the sample size of the independent study estimating $\boldsymbol{\beta}$. Then $\widehat{PAF}$ and $\widehat{PIF}$ converge in probability to PAF and PIF, respectively, and both $\sqrt{n}(\widehat{PAF} - PAF)$ and $\sqrt{n}(\widehat{PIF} - PIF)$ are asymptotically mean-zero multivariate normal.

Methods - Approximate Method

Suppose we only had the mean $\bar{\boldsymbol{X}} = (\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_k)$ and variance $\hat{\sigma}_{i,j} = Cov(X_i, X_j)$ of the exposure. This is often what is reported in publications, where *individual-level data is not available.*

Methods - Approximate Method

Suppose we only had the mean $\bar{\boldsymbol{X}} = (\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_k)$ and variance $\hat{\sigma}_{i,j} = Cov(X_i, X_j)$ of the exposure. This is often what is reported in publications, where *individual-level data is not available.*

We can use a second-order Taylor expansion for $\hat{\mu}_{\boldsymbol{n}}^{\mathrm{obs}}(\hat{\boldsymbol{\beta}})$ to derive a point estimate using *only the mean and variance*, leading to the following PAF estimator

$$\widehat{\mathrm{PAF}} = 1 - \frac{1}{RR(\bar{\boldsymbol{X}}; \hat{\boldsymbol{\beta}}) + \frac{1}{2}\sum_{i,j} \hat{\sigma}_{i,j} \frac{\partial^2 RR\left(\boldsymbol{X}, \hat{\boldsymbol{\beta}}\right)}{\partial X_i \partial X_j}\Big|_{\boldsymbol{X} = \bar{\boldsymbol{X}}}},$$

Methods - Approximate Method

Suppose we only had the mean $\bar{\boldsymbol{X}} = (\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_k)$ and variance $\hat{\sigma}_{i,j} = Cov(X_i, X_j)$ of the exposure. This is often what is reported in publications, where *individual-level data is not available.*

We can use a second-order Taylor expansion for $\hat{\mu}_n^{\mathrm{obs}}(\hat{\boldsymbol{\beta}})$ to derive a point estimate using *only the mean and variance*, leading to the following PAF estimator

$$\widehat{\mathrm{PAF}} = 1 - \frac{1}{RR(\bar{\boldsymbol{X}}; \hat{\boldsymbol{\beta}}) + \frac{1}{2}\sum_{i,j} \hat{\sigma}_{i,j} \frac{\partial^2 RR\left(\boldsymbol{X}, \hat{\boldsymbol{\beta}}\right)}{\partial X_i \partial X_j}\Big|_{\boldsymbol{X}=\bar{\boldsymbol{X}}}},$$

Repeat for $\hat{\mu}^{\mathrm{cft}}(\hat{\boldsymbol{\beta}})$ for the PIF.

Methods - Approximate Method

Suppose we only had the mean $\bar{\boldsymbol{X}} = (\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_k)$ and variance $\hat{\sigma}_{i,j} = Cov(X_i, X_j)$ of the exposure. This is often what is reported in publications, where *individual-level data is not available.*

We can use a second-order Taylor expansion for $\hat{\mu}_n^{\text{obs}}(\hat{\boldsymbol{\beta}})$ to derive a point estimate using *only the mean and variance*, leading to the following PAF estimator

$$\widehat{\text{PAF}} = 1 - \frac{1}{RR(\bar{\boldsymbol{X}}; \hat{\boldsymbol{\beta}}) + \frac{1}{2}\sum_{i,j} \hat{\sigma}_{i,j} \frac{\partial^2 RR\left(\boldsymbol{X}, \hat{\boldsymbol{\beta}}\right)}{\partial X_i \partial X_j}\Big|_{\boldsymbol{X}=\bar{\boldsymbol{X}}}},$$

Repeat for $\hat{\mu}^{\text{cft}}(\hat{\boldsymbol{\beta}})$ for the PIF.

To derive the variance, we apply the multivariate delta method, as the PIF and PAF are functions of three components: $\bar{\boldsymbol{X}}, \hat{\sigma}_{i,j}, \hat{\boldsymbol{\beta}}$

Simulation Studies

- Define true exposure as a mixture $p_0 + (1 - p_0)f(x)$, where $f(x)$ is a known parametric distribution, truncated at $M = 12$.
  Get true PAF value.
- For each simulation $b = 1, \ldots, B$, varying N:
  - Generate data from true underlying exposure distribution
  - Estimate the PAF and 95% confidence interval using the approximate and empirical methods
- Report coverage and average relative bias over the B simulations

Simulation Studies

True dist. $p_0 + (1 - p_0)f(x)$				Empirical		Approximate	
$f(x)$	p_0	true PAF	N	Rel. Bias %	Coverage %	Rel. Bias %	Coverage %
Lognormal	0.00	0.364	100				
			1000				
			10000				
	0.05	0.352	100				
			1000				
			10000				
	0.25	0.301	100				
			1000				
			10000				
	0.50	0.223	100				
			1000				
			10000				
	0.75	0.125	100				
			1000				
			10000				
Weibull	0.00	0.350	100				
			1000				
			10000				
	0.05	0.3385	100				
			1000				
			10000				
	0.25	0.288	100				
			1000				
			10000				
	0.50	0.212	100				
			1000				
			10000				
	0.75	0.119	100				
			1000				
			10000				

Simulation Studies

True dist. $p_0 + (1 - p_0) f(x)$				Empirical		Approximate	
$f(x)$	p_0	true PAF	N	Rel. Bias %	Coverage %	Rel. Bias %	Coverage %
Lognormal	0.00	0.364	100	-0.6	94.4	-0.8	94.5
			1000	-0.2	94.8	-0.5	95.0
			10000	-0.1	94.8	-0.3	94.8
	0.05	0.352	100	-0.6	95.5	-1.0	95.7
			1000	-0.2	94.8	-0.7	94.9
			10000	-0.1	94.7	-0.6	94.7
	0.25	0.301	100	-0.5	94.9	-2.2	94.8
			1000	0.0	95.0	-1.8	94.9
			10000	-0.1	95.1	-1.8	94.5
	0.50	0.223	100	-0.1	93.9	-3.0	93.2
			1000	0.0	95.0	-2.9	94.1
			10000	0.1	94.6	-2.8	93.9
	0.75	0.125	100	0.3	91.8	-1.3	91.1
			1000	0.1	94.6	-1.5	94.0
			10000	0.1	94.7	-1.5	94.4
Weibull	0.00	0.035	100	-0.5	93.9	-4.4	93.2
			1000	-0.1	94.8	-4.2	93.8
			10000	0.0	94.9	-4.1	92.5
	0.05	0.339	100	-0.5	95.2	-4.8	94.4
			1000	-0.1	94.8	-4.5	93.6
			10000	0.0	94.7	-4.4	91.9
	0.25	0.288	100	-0.4	94.3	-6.1	92.6
			1000	0.1	95.0	-5.7	92.8
			10000	0.0	95.0	-5.7	89.2
	0.50	0.212	100	-0.3	92.6	-7.0	89.8
			1000	0.2	94.7	-6.7	90.7
			10000	0.3	94.6	-6.6	88.0
	0.75	0.119	100	0.6	90.1	-4.7	88.2
			1000	0.4	94.7	-4.9	92.1
			10000	0.2	94.8	-4.8	91.7

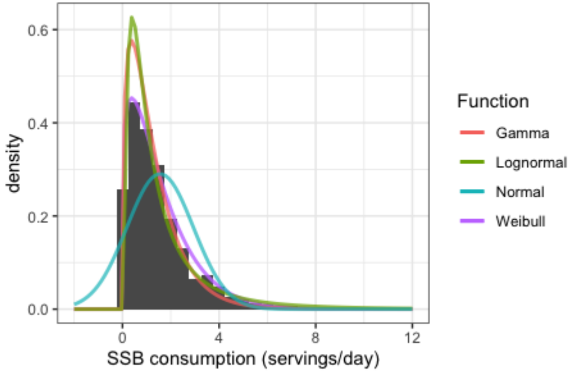
Illustrative Example

- Q: What proportion of type 2 diabetes cases can be attributed to sugar-sweetened beverage consumption in Mexico?
- SSB consumption data ($n = 7762$) from ENSANUT 2016[1]
- Meta-analytic relative risk taken from the Mexican Teacher's Cohort[2]

[1] Gaona-Pineda et al. 2018, [2] Stern et al. 2019

Illustrative Example

- Q: What proportion of type 2 diabetes cases can be attributed to sugar-sweetened beverage consumption in Mexico?
- SSB consumption data ($n = 7762$) from ENSANUT 2016[1]
- Meta-analytic relative risk taken from the Mexican Teacher's Cohort[2]



[1] Gaona-Pineda et al. 2018, [2] Stern et al. 2019

Illustrative Example

	Parameters	PAF (95% CI)
Standard Gamma	$k = 1.15, \theta = 1.29$	0.345
Mixture Gamma	$k = 1.41, \theta = 0.90$	0.280
Mixture Gamma ($M = 12$)	$k = 1.41, \theta = 0.90$	0.290
Standard Lognormal	$\log \mu = 0.082, \log \sigma = 0.31$	1
Mixture Lognormal	$\log \mu = 0.05, \log \sigma = 0.98$	1
Mixture Lognormal ($M = 12$)	$\log \mu = 0.05, \log \sigma = 0.98$	0.379
Standard Normal	$\mu = 1.48, \sigma = 1.38$	0.278
Mixture Normal	$\mu = 1.56, \sigma = 1.37$	0.375
Mixture Normal ($M = 12$)	$\mu = 1.56, \sigma = 1.37$	0.375
Standard Weibull	$k = 1.08, \lambda = 1.53$	$\boxed{0.345}$
Mixture Weibull	$k = 1.20, \lambda = 1.66$	0.339
Mixture Weibull ($M = 12$)	$k = 1.20, \lambda = 1.66$	0.339
Empirical	-	$\boxed{0.345}$ (0.224, 0.467)
Approximate	-	$\boxed{0.325}$ (0.219, 0.431)

Illustrative Example

	Parameters	PAF (95% CI)
Standard Gamma	$k = 1.15, \theta = 1.29$	0.345
Mixture Gamma	$k = 1.41, \theta = 0.90$	0.280
Mixture Gamma ($M = 12$)	$k = 1.41, \theta = 0.90$	0.290
Standard Lognormal	$\log \mu = 0.082, \log \sigma = 0.31$	1
Mixture Lognormal	$\log \mu = 0.05, \log \sigma = 0.98$	1
Mixture Lognormal ($M = 12$)	$\log \mu = 0.05, \log \sigma = 0.98$	0.379
Standard Normal	$\mu = 1.48, \sigma = 1.38$	0.278
Mixture Normal	$\mu = 1.56, \sigma = 1.37$	0.375
Mixture Normal ($M = 12$)	$\mu = 1.56, \sigma = 1.37$	0.375
Standard Weibull	$k = 1.08, \lambda = 1.53$	$\boxed{0.345}$
Mixture Weibull	$k = 1.20, \lambda = 1.66$	0.339
Mixture Weibull ($M = 12$)	$k = 1.20, \lambda = 1.66$	0.339
Empirical	-	$\boxed{0.345}$ (0.224, 0.467)
Approximate	-	$\boxed{0.325}$ (0.219, 0.431)

`pifpaf` `R` package available at
`https://github.com/colleenchan/pifpaf`

Recap

- PIF estimation requires an assumed distribution for the exposure
  - Biased when exposure distribution is misspecified and undefined when a heavily-tailed distribution is chosen
- We propose two nonparametric methods to estimate the PIF, both of which do not require making any distributional assumptions
  - Empirical method: Requires individual-level data
  - Approximate method: Requires only the mean and variance
- Conducted simulation studies of our methods
- PAF estimation of SSB consumption on type 2 diabetes incidence in Mexico (≈ 0.33)
- Possible extensions: nonparametric Bayesian inference, robust mean estimators for the PIF

Thank you!

Email: colleen.chan@yale.edu

The PIF and PAF as Causal Estimands

The PIF and PAF can interpreted as causal estimands if the following assumptions hold:

- β is a causal parameter, i.e., the model used to estimate relative risk adjusts for all known confounders
- β is transportable to the counterfactual population
- No effect modifiers

Methods - Approximate Method

Consider a general function $h(\boldsymbol{X})$, which is twice differentiable. Let

$$\boldsymbol{D}h(\boldsymbol{X}) = \frac{\partial h(\boldsymbol{X})}{\partial \boldsymbol{X}} \quad \text{and} \quad \boldsymbol{H}h(\boldsymbol{X}) = \frac{\partial^2 h(\boldsymbol{X})}{\partial \boldsymbol{X} \partial \boldsymbol{X}^T}.$$

The second-order Taylor polynomial for $h(\boldsymbol{X})$ is

$$\begin{aligned} h(\boldsymbol{X}) &\approx h(\widehat{\boldsymbol{\mu}}) + \boldsymbol{D}h(\widehat{\boldsymbol{\mu}})(\boldsymbol{X} - \widehat{\boldsymbol{\mu}}) + \frac{1}{2}(\boldsymbol{X} - \widehat{\boldsymbol{\mu}})^T \boldsymbol{H}h(\widehat{\boldsymbol{\mu}})(\boldsymbol{X} - \widehat{\boldsymbol{\mu}}) \\ &= h(\widehat{\boldsymbol{\mu}}) + \boldsymbol{D}h(\widehat{\boldsymbol{\mu}})(\boldsymbol{X} - \widehat{\boldsymbol{\mu}}) + \frac{1}{2} tr \left[(\boldsymbol{X} - \widehat{\boldsymbol{\mu}})(\boldsymbol{X} - \widehat{\boldsymbol{\mu}})^T \boldsymbol{H}h(\widehat{\boldsymbol{\mu}}) \right]. \end{aligned}$$

The first and second moments of $\boldsymbol{X}$ are

$$\boldsymbol{\mu}_{\boldsymbol{X}} = \mathbb{E}(\boldsymbol{X}) \quad \text{and} \quad \boldsymbol{\Sigma}_{\boldsymbol{X}} = Var(\boldsymbol{X}),$$

and their estimates are

$$\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}} = \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{X}_i) \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{X}} = \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{X}_i - \widehat{\boldsymbol{\mu}}_{\boldsymbol{X}})(\boldsymbol{X}_i - \widehat{\boldsymbol{\mu}}_{\boldsymbol{X}})^T.$$

Methods - Approximate Method

Applying the approximation to all subjects $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, ..., $\boldsymbol{X}_n$, we have

$$\frac{1}{n}\sum_{i=1}^{n} h(\boldsymbol{X}_i) \approx h(\widehat{\boldsymbol{\mu}}) + \frac{1}{2} tr\left[\widehat{\Sigma}_{\boldsymbol{X}} \boldsymbol{H} h(\widehat{\boldsymbol{\mu}})\right].$$

Using this, we can approximate the following scalar functions,

$$\hat{\mu}_n^{\mathrm{obs}}(\hat{\boldsymbol{\beta}}), \hat{\mu}_n^{\mathrm{cft}}(\hat{\boldsymbol{\beta}}), \frac{1}{n}\sum_{i=1}^{n}(RR(X_i;\hat{\boldsymbol{\beta}}))^2, \frac{1}{n}\sum_{i=1}^{n}(RR(g(X_i);\hat{\boldsymbol{\beta}}))^2, \frac{1}{n}\sum_{i=1}^{n} RR(X_i;\hat{\boldsymbol{\beta}})RR(g(X_i);\hat{\boldsymbol{\beta}}),$$

and the following vector functions, entry by entry,

$$\frac{1}{n}\sum_{i=1}^{n}\nabla_{\boldsymbol{\beta}} RR(X_i;\boldsymbol{\beta})\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} \text{ and } \frac{1}{n}\sum_{i=1}^{n}\nabla_{\boldsymbol{\beta}} RR(g(X_i);\boldsymbol{\beta})\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}}.$$

These calculations appear in the confidence intervals for $\widehat{\text{PAF}}$ and $\widehat{\text{PIF}}$.

pifpaf R package[1]

```
> x <- df$serving
> pif.ind(x, beta = log(1.27), varbeta = 0.002)
Estimating PAF and 95% confidence interval
$pif
[1] 0.3453349

$ci
[1] 0.2223366 0.4683332

> pif.app(meanx = 1.483, varx = 1.909, n = 7762, beta = log(1.27), varbeta = 0.002)
Estimating PAF and 95% confidence interval
$pif
[1] 0.3250859

$ci
[1] 0.2184401 0.4317316

> pif.ind(x, beta = log(1.27), varbeta = 0.002, a = -1)
Estimating PIF with counterfactual exposure g(x) = x - 1 and 95% confidence interval
$pif
[1] 0.2036881

$ci
[1] 0.08078843 0.32658777
```

[1] https://github.com/colleenchan/pifpaf