Package ‘powerMediation’

March 23, 2021

Version 0.3.4

Date 2021-03-22

Title Power/Sample Size Calculation for Mediation Analysis

Author Weiliang Qiu <weiliang.qiu@gmail.com>

Maintainer Weiliang Qiu <weiliang.qiu@gmail.com>

Depends R (>= 3.5.0), stats

Description Functions to
calculate power and sample size for testing
(1) mediation effects;
(2) the slope in a simple linear regression;
(3) odds ratio in a simple logistic regression;
(4) mean change for longitudinal study with 2 time points;
(5) interaction effect in 2-way ANOVA; and
(6) the slope in a simple Poisson regression.

License GPL (>= 2)

Repository CRAN

Date/Publication 2021-03-23 14:20:02 UTC

NeedsCompilation no

R topics documented:

`minEffect.SLR` *Minimum detectable slope*

Description

Calculate minimal detectable slope given sample size and power for simple linear regression.

Usage

```
minEffect.SLR(n,
              power,
              sigma.x,
              sigma.y,
              alpha = 0.05,
              verbose = TRUE)
```

Arguments

`n`	sample size.
`power`	power for testing if $\lambda = 0$ for the simple linear regression $y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
`sigma.x`	standard deviation of the predictor $sd(x) = \sigma_x$.

sigma.y	marginal standard deviation of the outcome $sd(y) = \sigma_y$. (not the conditional standard deviation $sd(y\|x)$)
alpha	type I error rate.
verbose	logical. TRUE means printing minimum absolute detectable effect; FALSE means not printing minimum absolute detectable effect.

Details

The test is for testing the null hypothesis $\lambda = 0$ versus the alternative hypothesis $\lambda \neq 0$ for the simple linear regressions:

$$y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Value

lambda.a	minimum absolute detectable effect.
res.uniroot	results of optimization to find the optimal minimum absolute detectable effect.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set alpha=0.10 to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Dupont, W.D. and Plummer, W.D.. Power and Sample Size Calculations for Studies Involving Linear Regression. *Controlled Clinical Trials*. 1998;19:589-601.

See Also

power.SLR, power.SLR.rho, ss.SLR, ss.SLR.rho.

Examples

```
minEffect.SLR(n = 100, power = 0.8, sigma.x = 0.2, sigma.y = 0.5,
  alpha = 0.05, verbose = TRUE)
```

`minEffect.VSMc` *Minimum detectable slope for mediator in linear regression based on Vittinghoff, Sen and McCulloch's (2009) method*

Description

Calculate minimal detectable slope for mediator given sample size and power in simple linear regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
minEffect.VSMc(n,
               power,
               sigma.m,
               sigma.e,
               corr.xm,
               alpha = 0.05,
               verbose = TRUE)
```

Arguments

`n`	sample size.
`power`	power for testing $b_2 = 0$ for the linear regression $y_i = b0 + b1x_i + b2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
`sigma.m`	standard deviation of the mediator.
`sigma.e`	standard deviation of the random error term in the linear regression $y_i = b0 + b1x_i + b2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
`corr.xm`	correlation between the predictor x and the mediator m.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing minimum absolute detectable effect; `FALSE` means not printing minimum absolute detectable effect.

Details

The test is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the linear regressions:

$$y_i = b_0 + b_1 x_i + b_2 m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Vittinghoff et al. (2009) showed that for the above linear regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$, if the correlation `corr.xm` between the primary predictor and mediator is non-zero.

The full model is

$$y_i = b_0 + b_1 x_i + b_2 m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

The reduced model is

$$y_i = b_0 + b_1 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`b2`	minimum absolute detectable effect.
`res.uniroot`	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu `<stwxq@channing.harvard.edu>`

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`powerMediation.VSMc`, `ssMediation.VSMc`

Examples

```
# example in section 3 (page 544) of Vittinghoff et al. (2009).
# minimum effect is =0.1
minEffect.VSMc(n = 863, power = 0.8, sigma.m = 1,
  sigma.e = 1, corr.xm = 0.3, alpha = 0.05, verbose = TRUE)
```

`minEffect.VSMc.cox` *Minimum detectable slope for mediator in cox regression based on Vittinghoff, Sen and McCulloch's (2009) method*

Description

Calculate minimal detectable slope for mediator given sample size and power in cox regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
minEffect.VSMc.cox(n,
                   power,
                   sigma.m,
                   psi,
                   corr.xm,
                   alpha = 0.05,
                   verbose = TRUE)
```

Arguments

`n`	sample size.
`power`	power for testing $b_2 = 0$ for the cox regression $\log(\lambda) = \log(\lambda_0) + b1x_i + b2m_i$, where λ is the hazard function and λ_0 is the baseline hazard function.
`sigma.m`	standard deviation of the mediator.
`psi`	the probability that an observation is uncensored, so that the number of event $d = n * psi$, where n is the sample size.
`corr.xm`	correlation between the predictor x and the mediator m.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing minimum absolute detectable effect; `FALSE` means not printing minimum absolute detectable effect.

Details

The test is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the cox regressions:

$$\log(\lambda) = \log(\lambda_0) + b_1 x_i + b_2 m_i$$

Vittinghoff et al. (2009) showed that for the above cox regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$, if the correlation `corr.xm` between the primary predictor and mediator is non-zero.

The full model is

$$\log(\lambda) = \log(\lambda_0) + b_1 x_i + b_2 m_i$$

The reduced model is

$$\log(\lambda) = \log(\lambda_0) + b_1 x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`b2`	minimum absolute detectable effect.
`res.uniroot`	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

powerMediation.VSMc.cox, ssMediation.VSMc.cox

Examples

```
# example in section 6 (page 547) of Vittinghoff et al. (2009).
# minimum effect is = log(1.5) = 0.4054651

minEffect.VSMc.cox(n = 1399, power = 0.7999916,
  sigma.m = sqrt(0.25 * (1 - 0.25)), psi = 0.2, corr.xm = 0.3,
  alpha = 0.05, verbose = TRUE)
```

`minEffect.VSMc.logistic` *Minimum detectable slope for mediator in logistic regression based on Vittinghoff, Sen and McCulloch's (2009) method*

Description

Calculate minimal detectable slope for mediator given sample size and power in logistic regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
minEffect.VSMc.logistic(n,
                        power,
                        sigma.m,
                        p,
                        corr.xm,
                        alpha = 0.05,
                        verbose = TRUE)
```

Arguments

n	sample size.
power	power for testing $b_2 = 0$ for the logistic regression $\log(p_i/(1-p_i)) = b0 + b1x_i + b2m_i$.
sigma.m	standard deviation of the mediator.
p	the marginal prevalence of the outcome.
corr.xm	correlation between the predictor x and the mediator m.
alpha	type I error rate.
verbose	logical. TRUE means printing minimum absolute detectable effect; FALSE means not printing minimum absolute detectable effect.

Details

The test is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the logistic regressions:

$$\log(p_i/(1-p_i)) = b_0 + b_1 x_i + b_2 m_i$$

Vittinghoff et al. (2009) showed that for the above logistic regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$, if the correlation `corr.xm` between the primary predictor and mediator is non-zero.

The full model is

$$\log(p_i/(1-p_i)) = b_0 + b_1 x_i + b_2 m_i$$

The reduced model is

$$\log(p_i/(1-p_i)) = b_0 + b_1 x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

b2	minimum absolute detectable effect.
res.uniroot	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`powerMediation.VSMc.logistic`, `ssMediation.VSMc.logistic`

Examples

```
# example in section 4 (page 545) of Vittinghoff et al. (2009).
# minimum effect is log(1.5)= 0.4054651

minEffect.VSMc.logistic(n = 255, power = 0.8, sigma.m = 1,
  p = 0.5, corr.xm = 0.5, alpha = 0.05, verbose = TRUE)
```

---

minEffect.VSMc.poisson

Minimum detectable slope for mediator in poisson regression based on Vittinghoff, Sen and McCulloch's (2009) method

---

Description

Calculate minimal detectable slope for mediator given sample size and power in poisson regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
minEffect.VSMc.poisson(n,
                       power,
                       sigma.m,
                       EY,
                       corr.xm,
                       alpha = 0.05,
                       verbose = TRUE)
```

Arguments

`n`	sample size.
`power`	power for testing $b_2 = 0$ for the poisson regression $\log(E(Y_i)) = b0 + b1x_i + b2m_i$.
`sigma.m`	standard deviation of the mediator.
`EY`	the marginal mean of the outcome
`corr.xm`	correlation between the predictor x and the mediator m.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing minimum absolute detectable effect; `FALSE` means not printing minimum absolute detectable effect.

Details

The test is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the poisson regressions:

$$\log(E(Y_i)) = b_0 + b_1 x_i + b_2 m_i$$

Vittinghoff et al. (2009) showed that for the above poisson regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$, if the correlation `corr.xm` between the primary predictor and mediator is non-zero.

The full model is

$$\log(E(Y_i)) = b_0 + b_1 x_i + b_2 m_i$$

The reduced model is

$$\log(E(Y_i)) = b_0 + b_1 x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`b2`	minimum absolute detectable effect.
`res.uniroot`	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu `<stwxq@channing.harvard.edu>`

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`powerMediation.VSMc.poisson`, `ssMediation.VSMc.poisson`

Examples

```
# example in section 5 (page 546) of Vittinghoff et al. (2009).
# minimum effect is = log(1.35) = 0.3001046
minEffect.VSMc.poisson(n = 1239, power = 0.7998578,
  sigma.m = sqrt(0.25 * (1 - 0.25)),
  EY = 0.5, corr.xm = 0.5,
  alpha = 0.05, verbose = TRUE)
```

power.SLR *Power for testing slope for simple linear regression*

Description

Calculate power for testing slope for simple linear regression.

Usage

```
power.SLR(n,
          lambda.a,
          sigma.x,
          sigma.y,
          alpha = 0.05,
          verbose = TRUE)
```

Arguments

n	sample size.
lambda.a	regression coefficient in the simple linear regression $y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
sigma.x	standard deviation of the predictor $sd(x)$.
sigma.y	marginal standard deviation of the outcome $sd(y)$. (not the marginal standard deviation $sd(y\|x)$)
alpha	type I error rate.
verbose	logical. `TRUE` means printing power; `FALSE` means not printing power.

Details

The power is for testing the null hypothesis $\lambda = 0$ versus the alternative hypothesis $\lambda \neq 0$ for the simple linear regressions:

$$y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Value

power	power for testing if $b_2 = 0$.
delta	$\lambda \sigma_x \sqrt{n} / \sqrt{\sigma_y^2 - (\lambda \sigma_x)^2}$.
s	$\sqrt{\sigma_y^2 - (\lambda \sigma_x)^2}$.
t.cr	$\Phi^{-1}(1 - \alpha/2)$, where Φ is the cumulative distribution function of the standard normal distribution.
rho	correlation between the predictor x and outcome $y = \lambda \sigma_x / \sigma_y$.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Dupont, W.D. and Plummer, W.D.. Power and Sample Size Calculations for Studies Involving Linear Regression. *Controlled Clinical Trials*. 1998;19:589-601.

See Also

`minEffect.SLR`, `power.SLR.rho`, `ss.SLR.rho`, `ss.SLR`.

Examples

```
power.SLR(n=100, lambda.a=0.8, sigma.x=0.2, sigma.y=0.5,
  alpha = 0.05, verbose = TRUE)
```

`power.SLR.rho` *Power for testing slope for simple linear regression*

Description

Calculate power for testing slope for simple linear regression.

Usage

```
power.SLR.rho(n,
              rho2,
              alpha = 0.05,
              verbose = TRUE)
```

Arguments

`n`	sample size.
`rho2`	square of the correlation between the outcome and the predictor.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing power; `FALSE` means not printing power.

Details

The power is for testing the null hypothesis $\lambda = 0$ versus the alternative hypothesis $\lambda \neq 0$ for the simple linear regressions:

$$y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Value

`power`	power for testing if $b_2 = 0$.
`delta`	$\sqrt{n}/\sqrt{1/\rho^2 - 1}$.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Dupont, W.D. and Plummer, W.D.. Power and Sample Size Calculations for Studies Involving Linear Regression. *Controlled Clinical Trials*. 1998;19:589-601.

See Also

`minEffect.SLR`, `power.SLR`, `ss.SLR.rho`, `ss.SLR`.

Examples

```
power.SLR.rho(n=100, rho2=0.6, alpha = 0.05, verbose = TRUE)
```

`powerInteract2by2` *Power Calculation for Interaction Effect in 2x2 Two-Way ANOVA Given Effect Sizes*

Description

Power calculation for interaction effect in 2x2 two-way ANOVA given effect sizes.

Usage

```
powerInteract2by2(n, tauBetaSigma, alpha = 0.05, nTests = 1, verbose = FALSE)
```

Arguments

`n`	integer. Number of subjects per group.
`tauBetaSigma`	Effect sizes $(\tau\beta)_{ij}/\sigma, i = 1, \ldots, a, j = 1, \ldots, b$, where $a = b = 2$ and σ is the standard deviation of random error. Rows are for factor 1 and columns are for factor 2. Note that $\sum_{i=1}^{a} (\tau\beta)_{ij} = \sum_{j=1}^{b} (\tau\beta)_{ij} = 0$. We can get $(\tau\beta)_{11} = \theta$, $(\tau\beta)_{12} = -\theta$, $(\tau\beta)_{21} = -\theta$, $(\tau\beta)_{22} = \theta$. So `tauBetaSigma=`$\theta/\sigma$
`alpha`	family-wise type I error rate.
`nTests`	integer. For high-throughput omics study, we perform two-way ANOVA for each of 'nTests' probes. We use Bonferroni correction to control for family-wise type I error rate. That is, for each probe, type I error rate would be `alpha/nTests`.
`verbose`	logical. Indicating if intermediate results should be printed out.

Details

We assume the following model:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk},$$

where $i = 1, \ldots, a, j = 1, \ldots, b, k = 1, \ldots, n$, $\sum_{i=1}^{a} \tau_i = 0$, $\sum_{j=1}^{b} \beta_j = 0$, $\sum_{i=1}^{a} (\tau\beta)_{ij} = 0$, $\sum_{j=1}^{b} (\tau\beta)_{ij} = 0$, and $\epsilon_{ijk} \overset{i.i.d}{\sim} N\left(0, \sigma^2\right)$.
The group means are

$$\mu_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij}, i = 1 \ldots, a, j = 1, \ldots, b.$$

Note that $\mu = \sum_{i=1}^{a} \sum_{j=1}^{b} \mu_{ij}/(ab)$, $\tau_i = \sum_{j=1}^{b} \mu_{ij}/b - \mu$, and $\beta_j = \sum_{i=1}^{a} \mu_{ij}/a - \mu$.
The null hypothesis H_0: all $(\tau\beta)_{ij}, i = 1, \ldots, a, j = 1, \ldots, b$ are equal to zero. The alternative hypothesis H_a: at least one $(\tau\beta)_{ij}$ is different from zero.
The F test statistic is

$$F = MS_{AB}/MS_E \overset{H_a}{\sim} F_{(a-1)(b-1), ab(n-1), ncp},$$

where `ncp` is the non-centrality parameter of the F test statistic:

$$ncp = n \sum_{i=1}^{a} \sum_{j=1}^{b} \left[\frac{(\tau\beta)_{ij}}{\sigma}\right]^2.$$

For the scenario $a = b = 2$, we have $(\tau\beta)_{11} = \theta$, $(\tau\beta)_{12} = -\theta$, $(\tau\beta)_{21} = -\theta$, $(\tau\beta)_{22} = \theta$. Hence, the non-centrality parameter can be simplified to

$$ncp = 4n\left(\frac{\theta}{\sigma}\right)^2.$$

The power for testing the null hypothesis H_0 versus the alternative hypothesis H_a is

$$power = Pr\left(F > F_0 | H_a\right),$$

where the rejection region boundary F_0 satisfies:

$$Pr\left(F > F_0 | H_0\right) = \alpha/nTests.$$

Value

A list with 5 elements:

`power`	the power of the two-way ANOVA test
`df1`	the first degree of freedom of the F test statistic (`df1=(a-1)(b-1)`)
`df2`	the second degree of freedom of the F test statistic (`df1=a*b(n-1)`)
`F0`	the rejection region boundary
`ncp`	the non-centrality parameter

Author(s)

Weiliang Qiu <weiliang.qiu@gmail.com>

References

Chow SC, Shao J, and Wang H. Sample size calculations in clinical research. 2nd edition. Chapman & Hall/CRC. 2008

Montgomery DC. Design and Analysis of Experiments. 8th edition. John Wiley & Sons. Inc.

Examples

```
n = 25
tauBetaSigma = 0.3

# power = 0.8437275
res2 = powerInteract2by2(n = n, tauBetaSigma = tauBetaSigma,
    alpha = 0.05, nTests = 1, verbose = TRUE)
```

`powerLogisticBin` *Calculating power for simple logistic regression with binary predictor*

Description

Calculating power for simple logistic regression with binary predictor.

Usage

```
powerLogisticBin(n,
                 p1,
                 p2,
                 B,
                 alpha = 0.05)
```

Arguments

n	total number of sample size.
p1	$pr(diseased\|X = 0)$, i.e. the event rate at $X = 0$ in logistic regression $logit(p) = a + bX$, where X is the binary predictor.
p2	$pr(diseased\|X = 1)$, the event rate at $X = 1$ in logistic regression $logit(p) = a + bX$, where X is the binary predictor.
B	$pr(X = 1)$, i.e. proportion of the sample with $X = 1$
alpha	Type I error rate.

Details

The logistic regression mode is

$$\log(p/(1-p)) = \beta_0 + \beta_1 X$$

where $p = prob(Y = 1)$, X is the binary predictor, $p_1 = pr(diseased|X = 0)$, $p_2 = pr(diseased|X = 1)$, $B = pr(X = 1)$, and $p = (1 - B)p_1 + Bp_2$. The sample size formula we used for testing if $\beta_1 = 0$, is Formula (2) in Hsieh et al. (1998):

$$n = (Z_{1-\alpha/2}[p(1-p)/B]^{1/2} + Z_{power}[p_1(1-p_1) + p_2(1-p_2)(1-B)/B]^{1/2})^2/[(p_1-p_2)^2(1-B)]$$

where n is the required total sample size and Z_u is the u-th percentile of the standard normal distribution.

Value

Estimated power.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Hsieh, FY, Bloch, DA, and Larsen, MD. A SIMPLE METHOD OF SAMPLE SIZE CALCULATION FOR LINEAR AND LOGISTIC REGRESSION. *Statistics in Medicine*. 1998; 17:1623-1634.

See Also

powerLogisticBin

Examples

```
## Example in Table I Design (Balanced design with high event rates)
## of Hsieh et al. (1998 )
## the power = 0.95
powerLogisticBin(n = 1281, p1 = 0.4, p2 = 0.5, B = 0.5, alpha = 0.05)
```

`powerLogisticCon` *Calculating power for simple logistic regression with continuous predictor*

Description

Calculating power for simple logistic regression with continuous predictor.

Usage

```
powerLogisticCon(n,
                 p1,
                 OR,
                 alpha = 0.05)
```

Arguments

n	total sample size.
p1	the event rate at the mean of the continuous predictor X in logistic regression $logit(p) = a + bX$.
OR	Expected odds ratio. $\log(OR)$ is the change in log odds for the difference between at the mean of X and at one SD above the mean.
alpha	Type I error rate.

Details

The logistic regression mode is

$$\log(p/(1-p)) = \beta_0 + \beta_1 X$$

where $p = prob(Y = 1)$, X is the continuous predictor, and $\log(OR)$ is the the change in log odds for the difference between at the mean of X and at one SD above the mean. The sample size formula we used for testing if $\beta_1 = 0$ or equivalently $OR = 1$, is Formula (1) in Hsieh et al. (1998):

$$n = (Z_{1-\alpha/2} + Z_{power})^2/[p_1(1-p_1)[log(OR)]^2]$$

where n is the required total sample size, OR is the odds ratio to be tested, p_1 is the event rate at the mean of the predictor X, and Z_u is the u-th percentile of the standard normal distribution.

Value

Estimated power.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu `<stwxq@channing.harvard.edu>`

References

Hsieh, FY, Bloch, DA, and Larsen, MD. A SIMPLE METHOD OF SAMPLE SIZE CALCULATION FOR LINEAR AND LOGISTIC REGRESSION. *Statistics in Medicine*. 1998; 17:1623-1634.

See Also

`SSizeLogisticCon`

Examples

```
## Example in Table II Design (Balanced design (1)) of Hsieh et al. (1998 )
## the power is 0.95
powerLogisticCon(n=317, p1=0.5, OR=exp(0.405), alpha=0.05)
```

---

powerLong — *Power calculation for longitudinal study with 2 time point*

Description

Power calculation for testing if mean changes for 2 groups are the same or not for longitudinal study with 2 time point.

Usage

```
powerLong(es,
          n,
          rho = 0.5,
          alpha = 0.05)
```

Arguments

`es`	effect size of the difference of mean change.
`n`	sample size per group.
`rho`	correlation coefficient between baseline and follow-up values within a treatment group.
`alpha`	Type I error rate.

Details

The power formula is based on Equation 8.31 on page 336 of Rosner (2006).

$$power = \Phi\left(-Z_{1-\alpha/2} + \frac{\delta\sqrt{n}}{\sigma_d\sqrt{2}}\right)$$

where $\sigma_d = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$, $\delta = |\mu_1 - \mu_2|$, μ_1 is the mean change over time t in group 1, μ_2 is the mean change over time t in group 2, σ_1^2 is the variance of baseline values within a treatment group, σ_2^2 is the variance of follow-up values within a treatment group, ρ is the correlation coefficient between baseline and follow-up values within a treatment group, and Z_u is the u-th percentile of the standard normal distribution.

We wish to test $\mu_1 = \mu_2$.

When $\sigma_1 = \sigma_2 = \sigma$, then formula reduces to

$$power = \Phi\left(-Z_{1-\alpha/2} + \frac{|d|\sqrt{n}}{2\sqrt{1-\rho}}\right)$$

where $d = \delta/\sigma$.

Value

power for testing for difference of mean changes.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Rosner, B. *Fundamentals of Biostatistics*. Sixth edition. Thomson Brooks/Cole. 2006.

See Also

`ssLong`, `ssLongFull`, `powerLongFull`.

Examples

```
# Example 8.34 on page 336 of Rosner (2006)
# power=0.75
powerLong(es=5/15, n=75, rho=0.7, alpha=0.05)
```

`powerLong.multiTime` *Power calculation for testing if mean changes for 2 groups are the same or not for longitudinal study with more than 2 time points*

Description

Power calculation for testing if mean changes for 2 groups are the same or not for longitudinal study with more than 2 time points.

Usage

```
powerLong.multiTime(es, m, nn, sx2, rho = 0.5, alpha = 0.05)
```

Arguments

`es`	effect size
`m`	number of subjects
`nn`	number of observations per subject
`sx2`	within subject variance
`rho`	within subject correlation
`alpha`	type I error rate

Details

We are interested in comparing the slopes of the 2 groups A and B:

$$\beta_{1A} = \beta_{1B}$$

where

$$Y_{ijA} = \beta_{0A} + \beta_{1A} x_{jA} + \epsilon_{ijA}, j = 1, \ldots, nn; i = 1, \ldots, m$$

and

$$Y_{ijB} = \beta_{0B} + \beta_{1B} x_{jB} + \epsilon_{ijB}, j = 1, \ldots, nn; i = 1, \ldots, m$$

The power calculation formula is (Equation on page 30 of Diggle et al. (1994)):

$$power = \Phi \left[-z_{1-\alpha} + \sqrt{\frac{mnns_x^2 es^2}{2(1-\rho)}} \right]$$

where $es = d/\sigma$, d is the meaninful differnce of interest, $sigma^2$ is the variance of the random error, ρ is the within-subject correlation, and s_x^2 is the within-subject variance.

Value

power

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Diggle PJ, Liang KY, and Zeger SL (1994). Analysis of Longitundinal Data. page 30. Clarendon Press, Oxford

See Also

`ssLong.multiTime`

Examples

```
# power=0.8
powerLong.multiTime(es=0.5/10, m=196, nn=3, sx2=4.22, rho = 0.5, alpha = 0.05)
```

`powerLongFull` *Power calculation for longitudinal study with 2 time point*

Description

Power calculation for testing if mean changes for 2 groups are the same or not for longitudinal study with 2 time point.

Usage

```
powerLongFull(delta,
              sigma1,
              sigma2,
              n,
              rho = 0.5,
              alpha = 0.05)
```

Arguments

`delta`	absolute difference of the mean changes between the two groups: $\delta = \|\mu_1 - \mu_2\|$ where μ_1 is the mean change over time t in group 1, μ_2 is the mean change over time t in group 2.
`sigma1`	the standard deviation of baseline values within a treatment group
`sigma2`	the standard deviation of follow-up values within a treatment group
`n`	sample size per group

rho	correlation coefficient between baseline and follow-up values within a treatment group.
alpha	Type I error rate.

Details

The power formula is based on Equation 8.31 on page 336 of Rosner (2006).

$$power = \Phi\left(-Z_{1-\alpha/2} + \frac{\delta\sqrt{n}}{\sigma_d\sqrt{2}}\right)$$

where $\sigma_d = \sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2$, $\delta = |\mu_1-\mu_2|$, μ_1 is the mean change over time t in group 1, μ_2 is the mean change over time t in group 2, σ_1^2 is the variance of baseline values within a treatment group, σ_2^2 is the variance of follow-up values within a treatment group, ρ is the correlation coefficient between baseline and follow-up values within a treatment group, and Z_u is the u-th percentile of the standard normal distribution.

We wish to test $\mu_1 = \mu_2$.

Value

power for testing for difference of mean changes.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Rosner, B. *Fundamentals of Biostatistics*. Sixth edition. Thomson Brooks/Cole. 2006.

See Also

`ssLong`, `ssLongFull`, `powerLong`.

Examples

```
# Example 8.33 on page 336 of Rosner (2006)
# power=0.80
powerLongFull(delta=5, sigma1=15, sigma2=15, n=85, rho=0.7, alpha=0.05)
```

`powerMediation.Sobel` *Power for testing mediation effect (Sobel's test)*

Description

Calculate power for testing mediation effect based on Sobel's test.

Usage

```
powerMediation.Sobel(n,
                    theta.1a,
                    lambda.a,
                    sigma.x,
                    sigma.m,
                    sigma.epsilon,
                    alpha = 0.05,
                    verbose = TRUE)
```

Arguments

`n`	sample size.
`theta.1a`	regression coefficient for the predictor in the linear regression linking the predictor x to the mediator m ($m_i = \theta_0 + \theta_{1a} x_i + e_i, e_i \sim N(0, \sigma_e^2)$).
`lambda.a`	regression coefficient for the mediator in the linear regression linking the predictor x and the mediator m to the outcome y ($y_i = \gamma + \lambda_a m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$).
`sigma.x`	standard deviation of the predictor.
`sigma.m`	standard deviation of the mediator.
`sigma.epsilon`	standard deviation of the random error term in the linear regression linking the predictor x and the mediator m to the outcome y ($y_i = \gamma + \lambda_a m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$).
`alpha`	type I error.
`verbose`	logical. `TRUE` means printing power; `FALSE` means not printing power.

Details

The power is for testing the null hypothesis $\theta_1 \lambda = 0$ versus the alternative hypothesis $\theta_{1a} \lambda_a \neq 0$ for the linear regressions:

$$m_i = \theta_0 + \theta_{1a} x_i + e_i, e_i \sim N(0, \sigma_e^2)$$

$$y_i = \gamma + \lambda_a m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$$

Test statistic is based on Sobel's (1982) test:

$$Z = \frac{\hat{\theta}_{1a} \hat{\lambda_a}}{\hat{\sigma}_{\theta_{1a} \lambda_a}}$$

where $\hat{\sigma}_{\theta_{1a}\lambda_a}$ is the estimated standard deviation of the estimate $\hat{\theta}_{1a}\hat{\lambda_a}$ using multivariate delta method:

$$\sigma_{\theta_{1a}\lambda_a} = \sqrt{\theta_{1a}^2\sigma_{\lambda_a}^2 + \lambda_a^2\sigma_{\theta_{1a}}^2}$$

and $\sigma_{\theta_{1a}}^2 = \sigma_e^2/(n\sigma_x^2)$ is the variance of the estimate $\hat{\theta}_{1a}$, and $\sigma_{\lambda_a}^2 = \sigma_\epsilon^2/(n\sigma_m^2(1-\rho_{mx}^2))$ is the variance of the estimate $\hat{\lambda_a}$, σ_m^2 is the variance of the mediator m_i.

From the linear regression $m_i = \theta_0 + \theta_{1a}x_i + e_i$, we have the relationship $\sigma_e^2 = \sigma_m^2(1-\rho_{mx}^2)$. Hence, we can simply the variance $\sigma_{\theta_{1a},\lambda_a}$ to

$$\sigma_{\theta_{1a}\lambda_a} = \sqrt{\theta_{1a}^2\frac{\sigma_\epsilon^2}{n\sigma_m^2(1-\rho_{mx}^2)} + \lambda_a^2\frac{\sigma_m^2(1-\rho_{mx}^2)}{n\sigma_x^2}}$$

Value

`power` power of the test for the parameter $\theta_{1a}\lambda_a$

`delta` $\theta_1\lambda/(sd(\hat{\theta}_{1a})sd(\hat{\lambda}_a))$

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu `<stwxq@channing.harvard.edu>`

References

Sobel, M. E. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*. 1982;13:290-312.

See Also

`ssMediation.Sobel`, `testMediation.Sobel`

Examples

```
powerMediation.Sobel(n=248, theta.1a=0.1701, lambda.a=0.1998,
  sigma.x=0.57, sigma.m=0.61, sigma.epsilon=0.2,
  alpha = 0.05, verbose = TRUE)
```

`powerMediation.VSMc` *Power for testing mediation effect in linear regression based on Vittinghoff, Sen and McCulloch's (2009) method*

Description

Calculate Power for testing mediation effect in linear regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
powerMediation.VSMc(n,
                    b2,
                    sigma.m,
                    sigma.e,
                    corr.xm,
                    alpha = 0.05,
                    verbose = TRUE)
```

Arguments

`n`	sample size.
`b2`	regression coefficient for the mediator m in the linear regression $y_i = b0 + b1x_i + b2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
`sigma.m`	standard deviation of the mediator.
`sigma.e`	standard deviation of the random error term in the linear regression $y_i = b0 + b1x_i + b2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
`corr.xm`	correlation between the predictor x and the mediator m.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing power; `FALSE` means not printing power.

Details

The power is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the linear regressions:

$$y_i = b_0 + b_1x_i + b_2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Vittinghoff et al. (2009) showed that for the above linear regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$.

The full model is

$$y_i = b_0 + b_1x_i + b_2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

The reduced model is

$$y_i = b_0 + b_1x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

power	power for testing if $b_2 = 0$.
delta	$b_2\sigma_m\sqrt{1-\rho_{xm}^2}/\sigma_e$, where σ_m is the standard deviation of the mediator m, ρ_{xm} is the correlation between the predictor x and the mediator m, and σ_e is the standard deviation of the random error term in the linear regression.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

minEffect.VSMc, ssMediation.VSMc

Examples

```
# example in section 3 (page 544) of Vittinghoff et al. (2009).
# power=0.8
powerMediation.VSMc(n = 863, b2 = 0.1, sigma.m = 1, sigma.e = 1,
  corr.xm = 0.3, alpha = 0.05, verbose = TRUE)
```

---

powerMediation.VSMc.cox

Power for testing mediation effect in cox regression based on Vittinghoff, Sen and McCulloch's (2009) method

---

Description

Calculate Power for testing mediation effect in cox regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
powerMediation.VSMc.cox(n,
                        b2,
                        sigma.m,
                        psi,
                        corr.xm,
                        alpha = 0.05,
                        verbose = TRUE)
```

Arguments

`n`	sample size.
`b2`	regression coefficient for the mediator m in the cox regression $\log(\lambda) = \log(\lambda_0) + b1x_i + b2m_i$, where λ is the hazard function and λ_0 is the baseline hazard function.
`sigma.m`	standard deviation of the mediator.
`psi`	the probability that an observation is uncensored, so that the number of event $d = n * psi$, where n is the sample size.
`corr.xm`	correlation between the predictor x and the mediator m.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing power; `FALSE` means not printing power.

Details

The power is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the cox regressions:

$$\log(\lambda) = \log(\lambda_0) + b_1 x_i + b_2 m_i$$

where λ is the hazard function and λ_0 is the baseline hazard function.

Vittinghoff et al. (2009) showed that for the above cox regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$.

The full model is

$$\log(\lambda) = \log(\lambda_0) + b_1 x_i + b_2 m_i$$

The reduced model is

$$\log(\lambda) = \log(\lambda_0) + b_1 x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`power`	power for testing if $b_2 = 0$.
`delta`	$b_2 \sigma_m \sqrt{(1 - \rho_{xm}^2) psi}$

, where σ_m is the standard deviation of the mediator m, ρ_{xm} is the correlation between the predictor x and the mediator m, and psi is the probability that an observation is uncensored, so that the number of event $d = n * psi$, where n is the sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`minEffect.VSMc.cox`, `ssMediation.VSMc.cox`

Examples

```
# example in section 6 (page 547) of Vittinghoff et al. (2009).
# power = 0.7999916
powerMediation.VSMc.cox(n = 1399, b2 = log(1.5),
  sigma.m = sqrt(0.25 * (1 - 0.25)), psi = 0.2, corr.xm = 0.3,
  alpha = 0.05, verbose = TRUE)
```

---

powerMediation.VSMc.logistic

Power for testing mediation effect in logistic regression based on Vittinghoff, Sen and McCulloch's (2009) method

---

Description

Calculate Power for testing mediation effect in logistic regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
powerMediation.VSMc.logistic(n,
                             b2,
                             sigma.m,
                             p,
                             corr.xm,
                             alpha = 0.05,
                             verbose = TRUE)
```

Arguments

`n`	sample size.
`b2`	regression coefficient for the mediator m in the logistic regression $\log(p_i/(1-p_i)) = b0 + b1x_i + b2m_i$.
`sigma.m`	standard deviation of the mediator.
`p`	the marginal prevalence of the outcome.
`corr.xm`	correlation between the predictor x and the mediator m.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing power; `FALSE` means not printing power.

Details

The power is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the logistic regressions:

$$\log(p_i/(1-p_i)) = b0 + b1x_i + b2m_i$$

Vittinghoff et al. (2009) showed that for the above logistic regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$.

The full model is

$$\log(p_i/(1-p_i)) = b_0 + b_1x_i + b_2m_i$$

The reduced model is

$$\log(p_i/(1-p_i)) = b_0 + b_1x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`power`	power for testing if $b_2 = 0$.
`delta`	$b_2\sigma_m\sqrt{(1-\rho_{xm}^2)p(1-p)}$

, where σ_m is the standard deviation of the mediator m, ρ_{xm} is the correlation between the predictor x and the mediator m, and p is the marginal prevalence of the outcome.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

minEffect.VSMc.logistic, ssMediation.VSMc.logistic

Examples

```
# example in section 4 (page 545) of Vittinghoff et al. (2009).
# power = 0.8005793
powerMediation.VSMc.logistic(n = 255, b2 = log(1.5), sigma.m = 1,
  p = 0.5, corr.xm = 0.5, alpha = 0.05, verbose = TRUE)
```

---

powerMediation.VSMc.poisson

Power for testing mediation effect in poisson regression based on Vittinghoff, Sen and McCulloch's (2009) method

---

Description

Calculate Power for testing mediation effect in poisson regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
powerMediation.VSMc.poisson(n,
                            b2,
                            sigma.m,
                            EY,
                            corr.xm,
                            alpha = 0.05,
                            verbose = TRUE)
```

Arguments

n	sample size.
b2	regression coefficient for the mediator m in the poisson regression $\log(E(Y_i)) = b0 + b1x_i + b2m_i$.
sigma.m	standard deviation of the mediator.
EY	the marginal mean of the outcome.
corr.xm	correlation between the predictor x and the mediator m.
alpha	type I error rate.
verbose	logical. TRUE means printing power; FALSE means not printing power.

Details

The power is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the poisson regressions:

$$\log(E(Y_i)) = b0 + b1x_i + b2m_i$$

Vittinghoff et al. (2009) showed that for the above poisson regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$.

The full model is

$$\log(E(Y_i)) = b_0 + b_1x_i + b_2m_i$$

The reduced model is

$$\log(E(Y_i)) = b_0 + b_1x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`power`	power for testing if $b_2 = 0$.
`delta`	$b_2\sigma_m\sqrt{(1-\rho_{xm}^2)EY}$

, where σ_m is the standard deviation of the mediator m, ρ_{xm} is the correlation between the predictor x and the mediator m, and EY is the marginal mean of the outcome.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <`stwxq@channing.harvard.edu`>

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`minEffect.VSMc.poisson`, `ssMediation.VSMc.poisson`

Examples

```
# example in section 5 (page 546) of Vittinghoff et al. (2009).
# power = 0.7998578
powerMediation.VSMc.poisson(n = 1239, b2 = log(1.35),
  sigma.m = sqrt(0.25 * (1 - 0.25)), EY = 0.5, corr.xm = 0.5,
  alpha = 0.05, verbose = TRUE)
```

powerPoisson	*Power calculation for simple Poisson regression*

Description

Power calculation for simple Poisson regression. Assume the predictor is normally distributed.

Usage

```
powerPoisson(
    beta0,
    beta1,
    mu.x1,
    sigma2.x1,
    mu.T = 1,
    phi = 1,
    alpha = 0.05,
    N = 50)
```

Arguments

beta0	intercept
beta1	slope
mu.x1	mean of the predictor
sigma2.x1	variance of the predictor
mu.T	mean exposure time
phi	a measure of over-dispersion
alpha	type I error rate
N	toal sample size

Details

The simple Poisson regression has the following form:

$$Pr(Y_i = y_i|\mu_i, t_i) = \exp(-\mu_i t_i)(\mu_i t_i)^{y_i}/(y_i!)$$

where

$$\mu_i = \exp(\beta_0 + \beta_1 x_{1i})$$

We are interested in testing the null hypothesis $\beta_1 = 0$ versus the alternative hypothesis $\beta_1 = \theta_1$. Assume x_1 is normally distributed with mean μ_{x_1} and variance $\sigma^2_{x_1}$. The sample size calculation formula derived by Signorini (1991) is

$$N = \phi \frac{\left[z_{1-\alpha/2}\sqrt{V\left(b_1|\beta_1=0\right)} + z_{power}\sqrt{V\left(b_1|\beta_1=\theta_1\right)}\right]^2}{\mu_T \exp(\beta_0)\theta_1^2}$$

where ϕ is the over-dispersion parameter ($= var(y_i)/mean(y_i)$), α is the type I error rate, b_1 is the estimate of the slope β_1, β_0 is the intercept, μ_T is the mean exposure time, z_a is the $100 * a$-th lower percentile of the standard normal distribution, and $V(b_1|\beta_1 = \theta)$ is the variance of the estimate b_1 given the true slope $\beta_1 = \theta$.

The variances are

$$V(b_1|\beta_1 = 0) = \frac{1}{\sigma_{x_1}^2}$$

and

$$V(b_1|\beta_1 = \theta_1) = \frac{1}{\sigma_{x_1}^2} \exp\left[-\left(\theta_1\mu_{x_1} + \theta_1^2\sigma_{x_1}^2/2\right)\right]$$

Value

power

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Signorini D.F. (1991). Sample size for Poisson regression. Biometrika. Vol.78. no.2, pp. 446-50

See Also

See Also as `sizePoisson`

Examples

```
# power = 0.8090542
print(powerPoisson(
    beta0 = 0.1,
    beta1 = 0.5,
    mu.x1 = 0,
    sigma2.x1 = 1,
    mu.T = 1,
    phi = 1,
    alpha = 0.05,
    N = 28))
```

sizePoisson	*Sample size calculation for simple Poisson regression*

Description

Sample size calculation for simple Poisson regression. Assume the predictor is normally distributed. Two-sided test is used.

Usage

```
sizePoisson(
    beta0,
    beta1,
    mu.x1,
    sigma2.x1,
    mu.T = 1,
    phi = 1,
    alpha = 0.05,
    power = 0.8)
```

Arguments

`beta0`	intercept
`beta1`	slope
`mu.x1`	mean of the predictor
`sigma2.x1`	variance of the predictor
`mu.T`	mean exposure time
`phi`	a measure of over-dispersion
`alpha`	type I error rate
`power`	power

Details

The simple Poisson regression has the following form:

$$Pr(Y_i = y_i|\mu_i, t_i) = \exp(-\mu_i t_i)(\mu_i t_i)^{y_i}/(y_i!)$$

where

$$\mu_i = \exp(\beta_0 + \beta_1 x_{1i})$$

We are interested in testing the null hypothesis $\beta_1 = 0$ versus the alternative hypothesis $\beta_1 = \theta_1$. Assume x_1 is normally distributed with mean μ_{x_1} and variance $\sigma^2_{x_1}$. The sample size calculation formula derived by Signorini (1991) is

$$N = \phi \frac{\left[z_{1-\alpha/2}\sqrt{V\left(b_1|\beta_1 = 0\right)} + z_{power}\sqrt{V\left(b_1|\beta_1 = \theta_1\right)}\right]^2}{\mu_T \exp(\beta_0)\theta_1^2}$$

where ϕ is the over-dispersion parameter ($= var(y_i)/mean(y_i)$), α is the type I error rate, b_1 is the estimate of the slope β_1, β_0 is the intercept, μ_T is the mean exposure time, z_a is the $100 * a$-th lower percentile of the standard normal distribution, and $V(b_1|\beta_1 = \theta)$ is the variance of the estimate b_1 given the true slope $\beta_1 = \theta$.

The variances are

$$V(b_1|\beta_1 = 0) = \frac{1}{\sigma_{x_1}^2}$$

and

$$V(b_1|\beta_1 = \theta_1) = \frac{1}{\sigma_{x_1}^2} \exp\left[-\left(\theta_1 \mu_{x_1} + \theta_1^2 \sigma_{x_1}^2/2\right)\right]$$

Value

total sample size

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Signorini D.F. (1991). Sample size for Poisson regression. Biometrika. Vol.78. no.2, pp. 446-50

See Also

See Also as powerPoisson

Examples

```
# sample size = 28
print(sizePoisson(
  beta0 = 0.1,
  beta1 = 0.5,
  mu.x1 = 0,
  sigma2.x1 = 1,
  mu.T = 1,
  phi = 1,
  alpha = 0.05,
  power = 0.8))
```

ss.SLR *Sample size for testing slope for simple linear regression*

Description

Calculate sample size for testing slope for simple linear regression.

Usage

```
ss.SLR(power,
       lambda.a,
       sigma.x,
       sigma.y,
       n.lower = 2.01,
       n.upper = 1e+30,
       alpha = 0.05,
       verbose = TRUE)
```

Arguments

power	power for testing if $\lambda = 0$ for the simple linear regression $y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
lambda.a	regression coefficient in the simple linear regression $y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
sigma.x	standard deviation of the predictor $sd(x)$.
sigma.y	marginal standard deviation of the outcome $sd(y)$. (not the marginal standard deviation $sd(y\|x)$)
n.lower	lower bound for the sample size.
n.upper	upper bound for the sample size.
alpha	type I error rate.
verbose	logical. `TRUE` means printing sample size; `FALSE` means not printing sample size.

Details

The test is for testing the null hypothesis $\lambda = 0$ versus the alternative hypothesis $\lambda \neq 0$ for the simple linear regressions:

$$y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Value

n	sample size.
res.uniroot	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Dupont, W.D. and Plummer, W.D.. Power and Sample Size Calculations for Studies Involving Linear Regression. *Controlled Clinical Trials*. 1998;19:589-601.

See Also

`minEffect.SLR`, `power.SLR`, `power.SLR.rho`, `ss.SLR.rho`.

Examples

```
ss.SLR(power=0.8, lambda.a=0.8, sigma.x=0.2, sigma.y=0.5,
  alpha = 0.05, verbose = TRUE)
```

`ss.SLR.rho` *Sample size for testing slope for simple linear regression based on R2*

Description

Calculate sample size for testing slope for simple linear regression based on R2.

Usage

```
ss.SLR.rho(power,
           rho2,
           n.lower = 2.01,
           n.upper = 1e+30,
           alpha = 0.05,
           verbose = TRUE)
```

Arguments

`power`	power.
`rho2`	square of the correlation between the outcome and the predictor.
`n.lower`	lower bound of the sample size.
`n.upper`	upper bound o the sample size.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing sample size; `FALSE` means not printing sample size.

Details

The test is for testing the null hypothesis $\lambda = 0$ versus the alternative hypothesis $\lambda \neq 0$ for the simple linear regressions:

$$y_i = \gamma + \lambda x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Value

n	sample size.
res.uniroot	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Dupont, W.D. and Plummer, W.D.. Power and Sample Size Calculations for Studies Involving Linear Regression. *Controlled Clinical Trials*. 1998;19:589-601.

See Also

`minEffect.SLR`, `power.SLR`, `power.SLR.rho`, `ss.SLR`.

Examples

```
ss.SLR.rho(power=0.8, rho2=0.6, alpha = 0.05, verbose = TRUE)
```

SSizeLogisticBin — *Calculating sample size for simple logistic regression with binary predictor*

Description

Calculating sample size for simple logistic regression with binary predictor.

Usage

```
SSizeLogisticBin(p1,
                 p2,
                 B,
                 alpha = 0.05,
                 power = 0.8)
```

Arguments

p1	$pr(diseased\|X = 0)$, i.e. the event rate at $X = 0$ in logistic regression $logit(p) = a + bX$, where X is the binary predictor.
p2	$pr(diseased\|X = 1)$, the event rate at $X = 1$ in logistic regression $logit(p) = a + bX$, where X is the binary predictor.
B	$pr(X = 1)$, i.e. proportion of the sample with $X = 1$
alpha	Type I error rate.
power	power for testing if the odds ratio is equal to one.

Details

The logistic regression mode is

$$\log(p/(1-p)) = \beta_0 + \beta_1 X$$

where $p = prob(Y = 1)$, X is the binary predictor, $p_1 = pr(diseased|X = 0)$, $p_2 = pr(diseased|X = 1)$, $B = pr(X = 1)$, and $p = (1 - B)p_1 + Bp_2$. The sample size formula we used for testing if $\beta_1 = 0$, is Formula (2) in Hsieh et al. (1998):

$$n = (Z_{1-\alpha/2}[p(1-p)/B]^{1/2} + Z_{power}[p_1(1-p_1) + p_2(1-p_2)(1-B)/B]^{1/2})^2/[(p_1-p_2)^2(1-B)]$$

where n is the required total sample size and Z_u is the u-th percentile of the standard normal distribution.

Value

total sample size required.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu `<stwxq@channing.harvard.edu>`

References

Hsieh, FY, Bloch, DA, and Larsen, MD. A SIMPLE METHOD OF SAMPLE SIZE CALCULATION FOR LINEAR AND LOGISTIC REGRESSION. *Statistics in Medicine*. 1998; 17:1623-1634.

See Also

powerLogisticBin

Examples

```
## Example in Table I Design (Balanced design with high event rates)
## of Hsieh et al. (1998 )
## the sample size is 1281
SSizeLogisticBin(p1 = 0.4, p2 = 0.5, B = 0.5, alpha = 0.05, power = 0.95)
```

`SSizeLogisticCon` *Calculating sample size for simple logistic regression with continuous predictor*

Description

Calculating sample size for simple logistic regression with continuous predictor.

Usage

```
SSizeLogisticCon(p1,
                 OR,
                 alpha = 0.05,
                 power = 0.8)
```

Arguments

`p1`	the event rate at the mean of the continuous predictor `X` in logistic regression $logit(p) = a + bX$,
`OR`	Expected odds ratio. $\log(OR)$ is the change in log odds for the difference between at the mean of X and at one SD above the mean.
`alpha`	Type I error rate.
`power`	power for testing if the odds ratio is equal to one.

Details

The logistic regression mode is

$$\log(p/(1-p)) = \beta_0 + \beta_1 X$$

where $p = prob(Y = 1)$, X is the continuous predictor, and $\log(OR)$ is the the change in log odds for the difference between at the mean of X and at one SD above the mean. The sample size formula we used for testing if $\beta_1 = 0$ or equivalently $OR = 1$, is Formula (1) in Hsieh et al. (1998):

$$n = (Z_{1-\alpha/2} + Z_{power})^2/[p_1(1-p_1)[log(OR)]^2]$$

where n is the required total sample size, OR is the odds ratio to be tested, p_1 is the event rate at the mean of the predictor X, and Z_u is the u-th percentile of the standard normal distribution.

Value

total sample size required.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Hsieh, FY, Bloch, DA, and Larsen, MD. A SIMPLE METHOD OF SAMPLE SIZE CALCULATION FOR LINEAR AND LOGISTIC REGRESSION. *Statistics in Medicine*. 1998; 17:1623-1634.

See Also

`powerLogisticCon`

Examples

```
## Example in Table II Design (Balanced design (1)) of Hsieh et al. (1998 )
## the sample size is 317
SSizeLogisticCon(p1 = 0.5, OR = exp(0.405), alpha = 0.05, power = 0.95)
```

---

ssLong *Sample size calculation for longitudinal study with 2 time point*

Description

Sample size calculation for testing if mean changes for 2 groups are the same or not for longitudinal study with 2 time point.

Usage

```
ssLong(es,
       rho = 0.5,
       alpha = 0.05,
       power = 0.8)
```

Arguments

`es`	effect size of the difference of mean change.
`rho`	correlation coefficient between baseline and follow-up values within a treatment group.
`alpha`	Type I error rate.
`power`	power for testing for difference of mean changes.

Details

The sample size formula is based on Equation 8.30 on page 335 of Rosner (2006).

$$n = \frac{2\sigma_d^2(Z_{1-\alpha/2} + Z_{power})^2}{\delta^2}$$

where $\sigma_d = \sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2$, $\delta = |\mu_1-\mu_2|$, μ_1 is the mean change over time t in group 1, μ_2 is the mean change over time t in group 2, σ_1^2 is the variance of baseline values within a treatment group, σ_2^2 is the variance of follow-up values within a treatment group, ρ is the correlation coefficient between baseline and follow-up values within a treatment group, and Z_u is the u-th percentile of the standard normal distribution.

We wish to test $\mu_1 = \mu_2$.

When $\sigma_1 = \sigma_2 = \sigma$, then formula reduces to

$$n = \frac{4(1-\rho)(Z_{1-\alpha/2} + Z_{\beta})^2}{d^2}$$

where $d = \delta/\sigma$.

Value

required sample size per group

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Rosner, B. *Fundamentals of Biostatistics*. Sixth edition. Thomson Brooks/Cole. 2006.

See Also

`ssLongFull`, `powerLong`, `powerLongFull`.

Examples

```
# Example 8.33 on page 336 of Rosner (2006)
# n=85
ssLong(es=5/15, rho=0.7, alpha=0.05, power=0.8)
```

`ssLong.multiTime` *Sample size calculation for testing if mean changes for 2 groups are the same or not for longitudinal study with more than 2 time points*

Description

Sample size calculation for testing if mean changes for 2 groups are the same or not for longitudinal study with more than 2 time points.

Usage

```
ssLong.multiTime(es, power, nn, sx2, rho = 0.5, alpha = 0.05)
```

Arguments

`es`	effect size
`power`	power
`nn`	number of observations per subject
`sx2`	within subject variance
`rho`	within subject correlation
`alpha`	type I error rate

Details

We are interested in comparing the slopes of the 2 groups A and B:

$$\beta_{1A} = \beta_{1B}$$

where

$$Y_{ijA} = \beta_{0A} + \beta_{1A} x_{jA} + \epsilon_{ijA}, j = 1, \ldots, nn; i = 1, \ldots, m$$

and

$$Y_{ijB} = \beta_{0B} + \beta_{1B} x_{jB} + \epsilon_{ijB}, j = 1, \ldots, nn; i = 1, \ldots, m$$

The sample size calculation formula is (Equation on page 30 of Diggle et al. (1994)):

$$m = \frac{2\left(Z_{1-\alpha} + z_{power}\right)^2 (1-\rho)}{nn s_x^2 es^2}$$

where $es = d/\sigma$, d is the meaninful differnce of interest, $sigma^2$ is the variance of the random error, ρ is the within-subject correlation, and s_x^2 is the within-subject variance.

Value

subject per group

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Diggle PJ, Liang KY, and Zeger SL (1994). Analysis of Longitundinal Data. page 30. Clarendon Press, Oxford

See Also

`powerLong.multiTime`

Examples

```
# subject per group = 196
ssLong.multiTime(es=0.5/10, power=0.8, nn=3, sx2=4.22, rho = 0.5, alpha=0.05)
```

`ssLongFull` *Sample size calculation for longitudinal study with 2 time point*

Description

Sample size calculation for testing if mean changes for 2 groups are the same or not for longitudinal study with 2 time point.

Usage

```
ssLongFull(delta,
           sigma1,
           sigma2,
           rho = 0.5,
           alpha = 0.05,
           power = 0.8)
```

Arguments

`delta`	absolute difference of the mean changes between the two groups: $\delta = \|\mu_1 - \mu_2\|$ where μ_1 is the mean change over time t in group 1, μ_2 is the mean change over time t in group 2.
`sigma1`	the standard deviation of baseline values within a treatment group
`sigma2`	the standard deviation of follow-up values within a treatment group

`rho`	correlation coefficient between baseline and follow-up values within a treatment group.
`alpha`	Type I error rate
`power`	power for testing for difference of mean changes.

Details

The sample size formula is based on Equation 8.30 on page 335 of Rosner (2006).

$$n = \frac{2\sigma_d^2 (Z_{1-\alpha/2} + Z_{power})^2}{\delta^2}$$

where $\sigma_d = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$, $\delta = |\mu_1 - \mu_2|$, μ_1 is the mean change over time t in group 1, μ_2 is the mean change over time t in group 2, σ_1^2 is the variance of baseline values within a treatment group, σ_2^2 is the variance of follow-up values within a treatment group, ρ is the correlation coefficient between baseline and follow-up values within a treatment group, and Z_u is the u-th percentile of the standard normal distribution.

We wish to test $\mu_1 = \mu_2$.

Value

required sample size per group

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu `<stwxq@channing.harvard.edu>`

References

Rosner, B. *Fundamentals of Biostatistics*. Sixth edition. Thomson Brooks/Cole. 2006.

See Also

`ssLong`, `powerLong`, `powerLongFull`.

Examples

```
# Example 8.33 on page 336 of Rosner (2006)
# n=85
ssLongFull(delta=5, sigma1=15, sigma2=15, rho=0.7, alpha=0.05, power=0.8)
```

`ssMediation.Sobel` *Sample size for testing mediation effectd (Sobel's test)*

Description

Calculate sample size for testing mediation effect based on Sobel's test.

Usage

```
ssMediation.Sobel(power,
                  theta.1a,
                  lambda.a,
                  sigma.x,
                  sigma.m,
                  sigma.epsilon,
                  n.lower = 1,
                  n.upper = 1e+30,
                  alpha = 0.05,
                  verbose = TRUE)
```

Arguments

`power`	power of the test.
`theta.1a`	regression coefficient for the predictor in the linear regression linking the predictor x to the mediator m ($m_i = \theta_0 + \theta_{1a}x_i + e_i, e_i \sim N(0, \sigma_e^2)$).
`lambda.a`	regression coefficient for the mediator in the linear regression linking the predictor x and the mediator m to the outcome y ($y_i = \gamma + \lambda_a m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$).
`sigma.x`	standard deviation of the predictor.
`sigma.m`	standard deviation of the mediator.
`sigma.epsilon`	standard deviation of the random error term in the linear regression linking the predictor x and the mediator m to the outcome y ($y_i = \gamma + \lambda_a m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$).
`n.lower`	lower bound of the sample size.
`n.upper`	upper bound of the sample size.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing power; `FALSE` means not printing power.

Details

The sample size is for testing the null hypothesis $\theta_1\lambda = 0$ versus the alternative hypothesis $\theta_{1a}\lambda_a \neq 0$ for the linear regressions:

$$m_i = \theta_0 + \theta_{1a}x_i + e_i, e_i \sim N(0, \sigma_e^2)$$

$$y_i = \gamma + \lambda_a m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$$

Test statistic is based on Sobel's (1982) test:

$$Z = \frac{\hat{\theta}_{1a}\hat{\lambda_a}}{\hat{\sigma}_{\theta_{1a}\lambda_a}}$$

where $\hat{\sigma}_{\theta_{1a}\lambda_a}$ is the estimated standard deviation of the estimate $\hat{\theta}_{1a}\hat{\lambda_a}$ using multivariate delta method:

$$\sigma_{\theta_{1a}\lambda_a} = \sqrt{\theta_{1a}^2\sigma_{\lambda_a}^2 + \lambda_a^2\sigma_{\theta_{1a}}^2}$$

and $\sigma_{\theta_{1a}}^2 = \sigma_e^2/(n\sigma_x^2)$ is the variance of the estimate $\hat{\theta}_{1a}$, and $\sigma_{\lambda_a}^2 = \sigma_\epsilon^2/(n\sigma_m^2(1-\rho_{mx}^2))$ is the variance of the estimate $\hat{\lambda_a}$, σ_m^2 is the variance of the mediator m_i.

From the linear regression $m_i = \theta_0 + \theta_{1a}x_i + e_i$, we have the relationship $\sigma_e^2 = \sigma_m^2(1-\rho_{mx}^2)$. Hence, we can simply the variance $\sigma_{\theta_{1a},\lambda_a}$ to

$$\sigma_{\theta_{1a}\lambda_a} = \sqrt{\theta_{1a}^2\frac{\sigma_\epsilon^2}{n\sigma_m^2(1-\rho_{mx}^2)} + \lambda_a^2\frac{\sigma_m^2(1-\rho_{mx}^2)}{n\sigma_x^2}}$$

Value

`n`	sample size.
`res.uniroot`	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Sobel, M. E. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*. 1982;13:290-312.

See Also

`powerMediation.Sobel`, `testMediation.Sobel`

Examples

```
ssMediation.Sobel(power=0.8, theta.1a=0.1701, lambda.a=0.1998,
  sigma.x=0.57, sigma.m=0.61, sigma.epsilon=0.2,
  alpha = 0.05, verbose = TRUE)
```

ssMediation.VSMc *Sample size for testing mediation effect in linear regression based on Vittinghoff, Sen and McCulloch's (2009) method*

Description

Calculate sample size for testing mediation effect in linear regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
ssMediation.VSMc(power,
                 b2,
                 sigma.m,
                 sigma.e,
                 corr.xm,
                 n.lower = 1,
                 n.upper = 1e+30,
                 alpha = 0.05,
                 verbose = TRUE)
```

Arguments

`power`	power for testing $b_2 = 0$ for the linear regression $y_i = b0+b1x_i+b2m_i+\epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
`b2`	regression coefficient for the mediator m in the linear regression $y_i = b0 + b1x_i + b2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
`sigma.m`	standard deviation of the mediator.
`sigma.e`	standard deviation of the random error term in the linear regression $y_i = b0 + b1x_i + b2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$.
`corr.xm`	correlation between the predictor x and the mediator m.
`n.lower`	lower bound for the sample size.
`n.upper`	upper bound for the sample size.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing sample size; `FALSE` means not printing sample size.

Details

The test is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the linear regressions:

$$y_i = b_0 + b_1x_i + b_2m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Vittinghoff et al. (2009) showed that for the above linear regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$.

The full model is

$$y_i = b_0 + b_1 x_i + b_2 m_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

The reduced model is

$$y_i = b_0 + b_1 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_e^2)$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`n`	sample size.
`res.uniroot`	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`minEffect.VSMc`, `powerMediation.VSMc`

Examples

```
# example in section 3 (page 544) of Vittinghoff et al. (2009).
# n=863
ssMediation.VSMc(power = 0.80, b2 = 0.1, sigma.m = 1, sigma.e = 1,
  corr.xm = 0.3, alpha = 0.05, verbose = TRUE)
```

`ssMediation.VSMc.cox` *Sample size for testing mediation effect in cox regression based on Vittinghoff, Sen and McCulloch's (2009) method*

Description

Calculate sample size for testing mediation effect in cox regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
ssMediation.VSMc.cox(power,
                     b2,
                     sigma.m,
                     psi,
                     corr.xm,
                     n.lower = 1,
                     n.upper = 1e+30,
                     alpha = 0.05,
                     verbose=TRUE)
```

Arguments

`power`	power for testing $b_2 = 0$ for the cox regression $\log(\lambda) = \log(\lambda_0) + b1x_i + b2m_i$, where λ is the hazard function and λ_0 is the baseline hazard function.
`b2`	regression coefficient for the mediator m in the cox regression $\log(\lambda) = \log(\lambda_0) + b1x_i + b2m_i$, where λ is the hazard function and λ_0 is the baseline hazard function.
`sigma.m`	standard deviation of the mediator.
`psi`	the probability that an observation is uncensored, so that the number of event $d = n * psi$, where n is the sample size.
`corr.xm`	correlation between the predictor x and the mediator m.
`n.lower`	lower bound for the sample size.
`n.upper`	upper bound for the sample size.
`alpha`	type I error rate.
`verbose`	logical. `TRUE` means printing sample size; `FALSE` means not printing sample size.

Details

The test is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the cox regressions:

$$\log(\lambda) = \log(\lambda_0) + b1x_i + b2m_i$$

Vittinghoff et al. (2009) showed that for the above cox regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$.

The full model is

$$\log(\lambda) = \log(\lambda_0) + b_1 x_i + b_2 m_i$$

The reduced model is

$$\log(\lambda) = \log(\lambda_0) + b_1 x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`n`	sample size.
`res.uniroot`	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu `<stwxq@channing.harvard.edu>`

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`minEffect.VSMc.cox`, `powerMediation.VSMc.cox`

Examples

```
# example in section 6 (page 547) of Vittinghoff et al. (2009).
# n = 1399
ssMediation.VSMc.cox(power = 0.7999916, b2 = log(1.5),
  sigma.m = sqrt(0.25 * (1 - 0.25)), psi = 0.2, corr.xm = 0.3,
  alpha = 0.05, verbose = TRUE)
```

ssMediation.VSMc.logistic

Sample size for testing mediation effect in logistic regression based on Vittinghoff, Sen and McCulloch's (2009) method

Description

Calculate sample size for testing mediation effect in logistic regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
ssMediation.VSMc.logistic(power,
                          b2,
                          sigma.m,
                          p,
                          corr.xm,
                          n.lower = 1,
                          n.upper = 1e+30,
                          alpha = 0.05,
                          verbose = TRUE)
```

Arguments

power	power for testing $b_2 = 0$ for the logistic regression $\log(p_i/(1-p_i)) = b0 + b1x_i + b2m_i$.
b2	regression coefficient for the mediator m in the logistic regression $\log(p_i/(1-p_i)) = b0 + b1x_i + b2m_i$.
sigma.m	standard deviation of the mediator.
p	the marginal prevalence of the outcome.
corr.xm	correlation between the predictor x and the mediator m.
n.lower	lower bound for the sample size.
n.upper	upper bound for the sample size.
alpha	type I error rate.
verbose	logical. `TRUE` means printing sample size; `FALSE` means not printing sample size.

Details

The test is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the logistic regressions:

$$\log(p_i/(1-p_i)) = b_0 + b_1x_i + b_2m_i$$

Vittinghoff et al. (2009) showed that for the above logistic regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$.

The full model is

$$\log(p_i/(1-p_i)) = b_0 + b_1 x_i + b_2 m_i$$

The reduced model is

$$\log(p_i/(1-p_i)) = b_0 + b_1 x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

`n`	sample size.
`res.uniroot`	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu `<stwxq@channing.harvard.edu>`

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`minEffect.VSMc.logistic`, `powerMediation.VSMc.logistic`

Examples

```
# example in section 4 (page 545) of Vittinghoff et al. (2009).
# n=255

ssMediation.VSMc.logistic(power = 0.80, b2 = log(1.5), sigma.m = 1, p = 0.5,
  corr.xm = 0.5, alpha = 0.05, verbose = TRUE)
```

ssMediation.VSMc.poisson
Sample size for testing mediation effect in poisson regression based on Vittinghoff, Sen and McCulloch's (2009) method

Description

Calculate sample size for testing mediation effect in poisson regression based on Vittinghoff, Sen and McCulloch's (2009) method.

Usage

```
ssMediation.VSMc.poisson(power,
                         b2,
                         sigma.m,
                         EY,
                         corr.xm,
                         n.lower = 1,
                         n.upper = 1e+30,
                         alpha = 0.05,
                         verbose = TRUE)
```

Arguments

power	power for testing $b_2 = 0$ for the poisson regression $\log(E(Y_i)) = b0 + b1x_i + b2m_i$.
b2	regression coefficient for the mediator m in the poisson regression $\log(E(Y_i)) = b0 + b1x_i + b2m_i$.
sigma.m	standard deviation of the mediator.
EY	the marginal mean of the outcome.
corr.xm	correlation between the predictor x and the mediator m.
n.lower	lower bound for the sample size.
n.upper	upper bound for the sample size.
alpha	type I error rate.
verbose	logical. `TRUE` means printing sample size; `FALSE` means not printing sample size.

Details

The test is for testing the null hypothesis $b_2 = 0$ versus the alternative hypothesis $b_2 \neq 0$ for the poisson regressions:

$$\log(E(Y_i)) = b_0 + b_1 x_i + b_2 m_i$$

Vittinghoff et al. (2009) showed that for the above poisson regression, testing the mediation effect is equivalent to testing the null hypothesis $H_0 : b_2 = 0$ versus the alternative hypothesis $H_a : b_2 \neq 0$.

The full model is

$$\log(E(Y_i)) = b_0 + b_1 x_i + b_2 m_i$$

The reduced model is

$$\log(E(Y_i)) = b_0 + b_1 x_i$$

Vittinghoff et al. (2009) mentioned that if confounders need to be included in both the full and reduced models, the sample size/power calculation formula could be accommodated by redefining `corr.xm` as the multiple correlation of the mediator with the confounders as well as the predictor.

Value

n	sample size.
res.uniroot	results of optimization to find the optimal sample size.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Vittinghoff, E. and Sen, S. and McCulloch, C.E.. Sample size calculations for evaluating mediation. *Statistics In Medicine*. 2009;28:541-557.

See Also

`minEffect.VSMc.poisson`, `powerMediation.VSMc.poisson`

Examples

```
# example in section 5 (page 546) of Vittinghoff et al. (2009).
# n = 1239
ssMediation.VSMc.poisson(power = 0.7998578, b2 = log(1.35),
  sigma.m = sqrt(0.25 * (1 - 0.25)), EY = 0.5, corr.xm = 0.5,
  alpha = 0.05, verbose = TRUE)
```

testMediation.Sobel *P-value and confidence interval for testing mediation effect (Sobel's test)*

Description

Calculate p-value and confidence interval for testing mediation effect based on Sobel's test.

Usage

```
testMediation.Sobel(theta.1.hat,
                    lambda.hat,
                    sigma.theta1,
                    sigma.lambda,
                    alpha = 0.05)
```

Arguments

theta.1.hat	estimated regression coefficient for the predictor in the linear regression linking the predictor x to the mediator m ($m_i = \theta_0 + \theta_1 x_i + e_i, e_i \sim N(0, \sigma_e^2)$).
lambda.hat	estimated regression coefficient for the mediator in the linear regression linking the predictor x and the mediator m to the outcome y ($y_i = \gamma + \lambda m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$).
sigma.theta1	standard deviation of $\hat{\theta}_1$ in the linear regression linking the predictor x to the mediator m ($m_i = \theta_0 + \theta_1 x_i + e_i, e_i \sim N(0, \sigma_e^2)$).
sigma.lambda	standard deviation of $\hat{\lambda}$ in the linear regression linking the predictor x and the mediator m to the outcome y ($y_i = \gamma + \lambda m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$).
alpha	significance level of a test.

Details

The test is for testing the null hypothesis $\theta_1 \lambda = 0$ versus the alternative hypothesis $\theta_{1a}\lambda_a \neq 0$ for the linear regressions:

$$m_i = \theta_0 + \theta_1 x_i + e_i, e_i \sim N(0, \sigma_e^2)$$
$$y_i = \gamma + \lambda m_i + \lambda_2 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma_\epsilon^2)$$

Test statistic is based on Sobel's (1982) test:

$$Z = \frac{\hat{\theta}_1 \hat{\lambda}}{\hat{\sigma}_{\theta_1 \lambda}}$$

where $\hat{\sigma}_{\theta_1 \lambda}$ is the estimated standard deviation of the estimate $\hat{\theta}_1 \hat{\lambda}$ using multivariate delta method:

$$\sigma_{\theta_1 \lambda} = \sqrt{\theta_1^2 \sigma_\lambda^2 + \lambda^2 \sigma_{\theta_1}^2}$$

and $\hat{\sigma}_{\theta_1}$ is the estimated standard deviation of the estimate $\hat{\theta}_1$, and $\hat{\sigma}_\lambda$ is the estimated standard deviation of the estimate $\hat{\lambda}$.

Value

- `pval` p-value for testing the null hypothesis $\theta_1\lambda = 0$ versus the alternative hypothesis $\theta_{1a}\lambda_a \neq 0$.
- `CI.low` Lower bound of the $100(1-\alpha)\%$ confidence interval for the parameter $\theta_1\lambda$.
- `CI.upp` Upper bound of the $100(1-\alpha)\%$ confidence interval for the parameter $\theta_1\lambda$.

Note

The test is a two-sided test. For one-sided tests, please double the significance level. For example, you can set `alpha=0.10` to obtain one-sided test at 5% significance level.

Author(s)

Weiliang Qiu <stwxq@channing.harvard.edu>

References

Sobel, M. E. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*. 1982;13:290-312.

See Also

powerMediation.Sobel, ssMediation.Sobel

Examples

```
testMediation.Sobel(theta.1.hat=0.1701, lambda.hat=0.1998,
  sigma.theta1=0.01, sigma.lambda=0.02, alpha=0.05)
```

Index