Improved Inference for the Signal Significance

Igor Volobouev*, Alex Trindade†

*Department of Physics & Astronomy, Texas Tech University
†Department of Mathematics & Statistics, Texas Tech University

i.volobouev@ttu.edu, alex.trindade@ttu.edu

PHYSTAT Dark Matter 2019, Aug 1

Signal Search Models

Two statistical models are often encountered in signal searches:

- Signal/background density mixture

$$p(\boldsymbol{x}|\alpha) = \alpha s(\boldsymbol{x}) + (1-\alpha) b(\boldsymbol{x}). \tag{1}$$

  Parameter of interest is the signal fraction α.

- Poison process with the intensity function given by

$$\Lambda(\boldsymbol{x}|\lambda) = \lambda s(\boldsymbol{x}) + \mu b(\boldsymbol{x}). \tag{2}$$

  Parameter of interest is the signal rate λ. For the purposes of this study, it is assumed that the background rate μ is known.

$s(\boldsymbol{x})$ and $b(\boldsymbol{x})$ are continuous, normalized, and fully specified (i.e., no nuisance parameters) densities for the signal and background, respectively.

Signal Strength Determination by Maximum Likelihood

- The signal fraction α of the mixture model is estimated by maximizing the log-likelihood:

$$\ell(\alpha) = \sum_{i=1}^{N} \ln p(\boldsymbol{x}_i|\alpha). \tag{3}$$

$$\hat{\alpha} := \operatorname*{argmax}_{\alpha} \ell(\alpha) \tag{4}$$

- The signal rate λ of the Poisson process model is estimated by maximizing the (extended) log-likelihood:

$$\ell(\lambda) = -(\lambda + \mu) + \sum_{i=1}^{N} \ln \Lambda(\boldsymbol{x}_i|\lambda) \tag{5}$$

$$\hat{\lambda} := \operatorname*{argmax}_{\lambda} \ell(\lambda) \tag{6}$$

$\hat{\alpha}$ and $\hat{\lambda}$ are asymptotically consistent, asymptotically normal, and asymptotically efficient under some mild conditions on $s(\boldsymbol{x})$ and $b(\boldsymbol{x})$.

Testing for Signal Presence

- Wald test

$$T_W := \frac{\hat{\alpha}^2}{\sigma_{CR}^2}, \quad \sigma_{CR} = \left(\mathrm{E}\left[-\frac{d^2 \ell(\alpha)}{d\alpha^2} \right] \Bigg|_{\alpha=0} \right)^{-1/2} \tag{7}$$

- Alternative Wald test

$$T_{W2} := \hat{\alpha}^2 \left(-\frac{d^2 \ell(\alpha)}{d\alpha^2} \Bigg|_{\alpha=\hat{\alpha}} \right) \tag{8}$$

- Yet another Wald test, using σ_M obtained from the $\Delta\ell(\alpha) = -1/2$ rule

$$T_{W3} := \frac{\hat{\alpha}^2}{\sigma_M^2} \tag{9}$$

- Score test

$$T_S := \sigma_{CR}^2 \left(\frac{d\ell(\alpha)}{d\alpha} \Bigg|_{\alpha=0} \right)^2 \tag{10}$$

- Likelihood ratio test

$$T_{LR} := 2[\ell(\hat{\alpha}) - \ell(0)] \tag{11}$$

All of these statistics are χ_1^2 distributed in the limit $N \to \infty$. Tests for the λ of the Poisson process model are similar.

Testing for Signal Presence (cont'd)

- For finite samples, better tests can be constructed by using statistics which take into account the sign of $\hat{\alpha}$.
- One option is to use truncated versions of the statistics presented on the previous slide. The CCGV paper, for example, is using the truncated likelihood ratio statistic, T_{TLR}, which is $\frac{1}{2}\delta(0) + \frac{1}{2}\chi^2_1$ distributed as $N \to \infty$:

$$T_{TLR} := \begin{cases} 2[\ell(\hat{\alpha}) - \ell(0)] & \text{if } \hat{\alpha} \geq 0, \\ 0 & \text{if } \hat{\alpha} < 0 \end{cases} \tag{12}$$

- Another option, favored in the modern statistical literature, is to construct the *signed square root* versions of these statistics:

$$R = \text{sgn}(\hat{\alpha})\sqrt{T} \tag{13}$$

The R statistics are $\mathcal{N}(0, 1)$ distributed as $N \to \infty$.

- T_{LR}, in particular, gives rise to the signed likelihood ratio, R_{LR}:

$$R_{LR} := \text{sgn}(\hat{\alpha})\sqrt{2[\ell(\hat{\alpha}) - \ell(0)]} \tag{14}$$

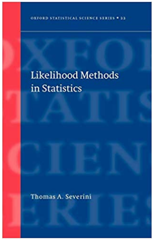
Nuisance Parameters in Searches

- We assume that all nuisance parameters (e.g., signal location and width) are treated by the theory of random fields. In application to particle physics searches, this technique is known as the Gross-Vitells method (GVM). The models and test statistics presented on the preceding slides correspond to a single point in the nuisance parameter space.
- The LEE factor determination by the GVM relies substantially on the assumptions of consistency and normality (or distribution according to χ_1^2) of the test statistic used. However, in order to be confident in the statements made for finite N, we must have an idea how well these assumptions are satisfied.
- Nuisance parameters in $b(\boldsymbol{x})$ can also be treated by profiling, but in this case mathematical derivations of improved asymptotic formulae become substantially more challenging.

Tools for Developing Higher-Order Asymptotic Theory

- Taylor expansion of $\ell(\alpha)$ near $\alpha = 0$ (or near "true value of the parameter" for other types of likelihoods)
- Joint cumulants for the derivatives of $\ell(\alpha)$
- Edgeworth series approximation of distributions

Some of the mathematical details are given in the backup slides. The derivations for $\ell(\alpha)$ follow the general methods discussed in the book by T.A. Severini, "Likelihood Methods in Statistics", Oxford University Press, 2000. Warning: a few of Severini's higher-order asymptotic formulae are incorrect. The formulae used in this study were rederived from first principles using Wolfram Mathematica.



Additional details about the mixture model simulations and asymptotic derivations are available in the article by I. Volobouev and A. Trindade, "Improved inference for the signal significance", 2018 JINST **13** P12011. The Poisson process model results and the $\mathcal{O}(N^{-5/2})$ approximations presented in this talk are new.

Gram-Charlier Expansion

- For a quantity whose distribution is close to the standard normal, an approximate density can be constructed with the Gram-Charlier expansion
$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \left(1 + \sum_{k=1}^{\infty} \beta_k H_k(z) \right). \tag{15}$$
The coefficients β_k are chosen to match the cumulants of the approximated distribution. In practice, the expansion is truncated at some maximum k.
- The CDF for this density is easily found using the following property of Hermite polynomials: $\int_{-\infty}^{z} H_k(\tau) e^{-\frac{\tau^2}{2}} d\tau = -H_{k-1}(z) e^{-\frac{z^2}{2}}$.
- A slightly different expansion can be obtained in the form
$$p(z) = \frac{1}{\sqrt{2\pi\kappa_2}} e^{-\frac{(z-\kappa_1)^2}{2\kappa_2}} \left(1 + \sum_{k=3}^{\infty} \beta'_k H_k\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}} \right) \right). \tag{16}$$

Edgeworth Expansion

- The Edgeworth expansion is obtained from the Gram-Charlier expansion by grouping together the terms with the same powers of $N^{-1/2}$.
- This grouping has to assume a particular asymptotic behavior of the cumulants. The typical cumulant behavior of the likelihood-based signed test statistics is $\kappa_1 \sim N^{-1/2}$, $\kappa_2 = 1 + \mathcal{O}(N^{-1})$, $\kappa_k \sim N^{-(k-2)/2}$ for all $k > 2$. To create, let say, $\mathcal{O}(N^{-2})$ expansion, cumulants up to and including κ_5 must be taken into account.
- The cumulants of R_{LR} are special: $\kappa_1 \sim N^{-1/2}$, $\kappa_2 = 1 + \mathcal{O}(N^{-1})$, $\kappa_k \sim N^{-k/2}$ for all $k > 2$ (this property of R_{LR} was established in 1999 by Mykland). Therefore, higher R_{LR} cumulants decay faster as $N \to \infty$ by factor $\sim \frac{1}{N}$. To create $\mathcal{O}(N^{-2})$ expansion, cumulants up to and including κ_3 must be taken into account.
- In combination with two different representations of Gram-Charlier series, this leads to four types of Edgeworth expansions to consider (the formulae can be found in the backup slides).

Cumulant Matching

- While the Edgeworth expansion is a useful asymptotic tool, it has some drawbacks. It is not guaranteed that the series are non-negative everywhere, and higher order cumulants of truncated series are not well-controlled.
- In 1939, Marcinkiewicz showed that the normal distribution is the only distribution whose cumulant generating function is a polynomial, i.e., the only distribution having a finite number of non-zero cumulants.
- This naturally leads to the question: if we know some number m of the leading cumulants, what is the distribution which minimizes, in some sense, the norms of all remaining cumulants? The minimization target can be formulated in a number of ways, something like $\sum_{j=m+1}^{\infty} |\kappa_j|^p$, $p > 0$ seems reasonable. As far as we know, this problem is unsolved.

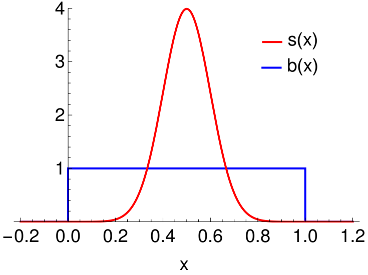
Shape Parameters

- For the mixture model, define
$V_k := \frac{1}{N} \, \mathrm{E}\left[\frac{d^k \ln \ell(\alpha)}{d\alpha^k}\right]\Big|_{\alpha=0} = \mathrm{E}\left[\frac{d^k \ln p(x|\alpha)}{d\alpha^k}\right]\Big|_{\alpha=0}$.
- For the Poisson process model, we observe only one realization of $\{N; \boldsymbol{x}_1, ..., \boldsymbol{x}_N\}$. It becomes natural to define
$V_k := \frac{1}{\mu} \, \mathrm{E}\left[\frac{d^k \ln \ell(\lambda)}{d\lambda^k}\right]\Big|_{\lambda=0}$ (i.e., V_k is defined "per unit μ").
- For the mixture and Poisson process models, it is convenient to represent higher-order asymptotic results in terms of the following parameters:

$$\gamma := \frac{V_3}{2 \, (-V_2)^{3/2}}, \; \rho := -\frac{V_4}{6 V_2^2}, \; \xi := \frac{V_5}{24 \, (-V_2)^{5/2}}, \; \zeta := \frac{V_6}{120 V_2^3} \quad (17)$$

$\left(\text{in general, } \frac{(-1)^{k-1} V_k}{(k-1)! \, (-V_2)^{k/2}}\right)$. These quantities are dimensionless and location-scale invariant. Large magnitudes of these parameters lead to slower convergence of the log-likelihood Taylor series and to stronger deviations from $\mathcal{N}(0, 1)$ asymptotic behavior.

An Example



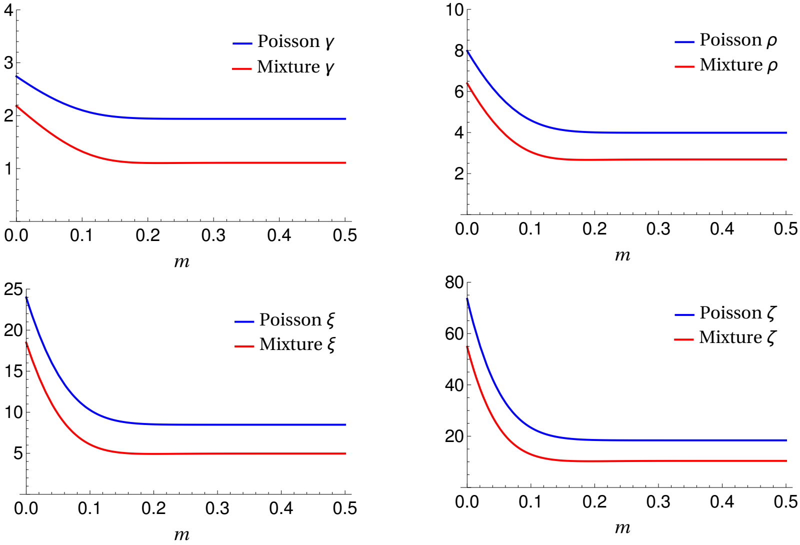
Use $b(x)$ uniform on $[0, 1]$ and truncated Gaussian $s(x)$:

$$b(x) = \begin{cases} 1 & \text{if } x \in [0,1] \\ 0 & \text{if } x \notin [0,1] \end{cases} , \tag{18}$$

$$s(x) = \begin{cases} e^{-\frac{(x-m)^2}{2\sigma^2}} \Big/ \int_0^1 e^{-\frac{(y-m)^2}{2\sigma^2}} dy & \text{if } x \in [0,1] \\ 0 & \text{if } x \notin [0,1] \end{cases} . \tag{19}$$

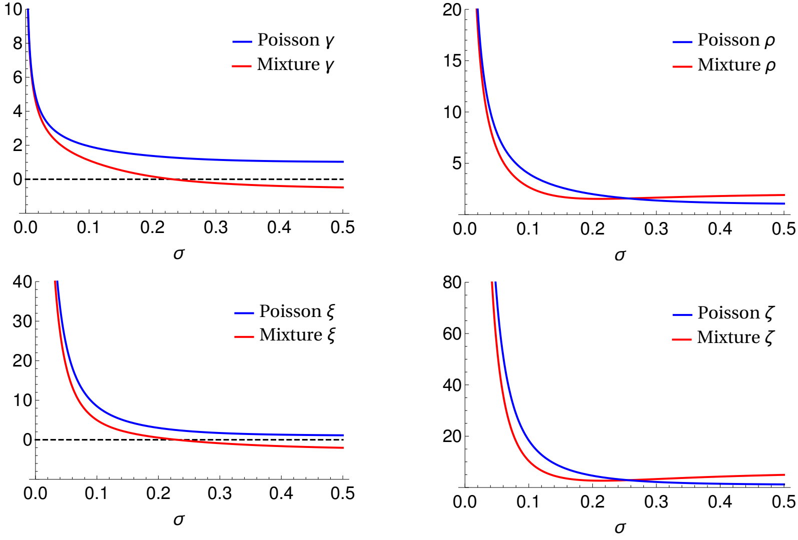
As the shape parameters are location-scale invariant, this example is applicable without loss of generality to a uniform background density and a truncated Gaussian signal density supported on an arbitrary interval $[a, b]$.

Dependence of Shape Parameters on Signal Location



For all of these plots, $\sigma = 0.1$.

Dependence of Shape Parameters on Signal Width



For all of these plots, $m = 0.5$.

Example Working Point

- Working point chosen for the example: $m = 0.5$ and $\sigma = 0.1$. For this point, the shape parameter values are as follows:

	Mixture	Poisson
γ	1.1094	1.9394
ρ	2.6893	3.9894
ξ	4.9671	8.4755
ζ	10.369	18.378

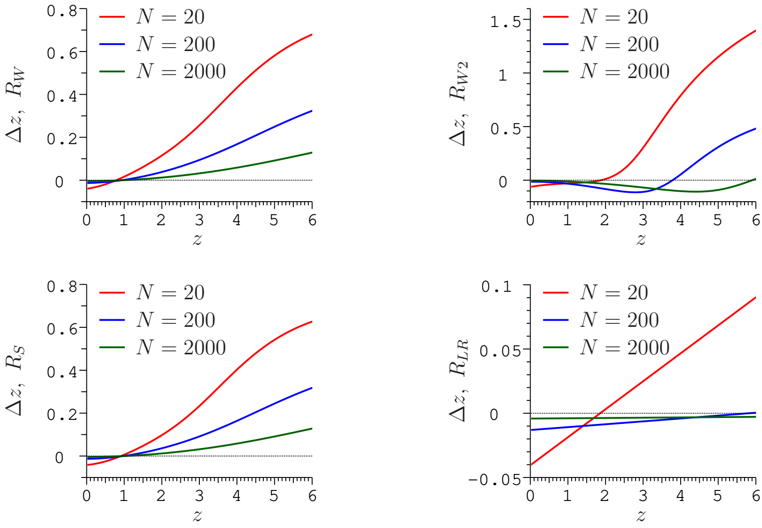
- For the mixture model, $\sigma_{CR} = \frac{1}{\sqrt{-NV_2}} \approx \frac{0.741}{\sqrt{N}}$. For $N \lessapprox 14$, it becomes impossible to distinguish $\alpha = 1$ (pure signal) from $\alpha = 0$ (pure background) at the level of 5σ.

- For the Poisson process model, $V_2/V_{2,C} \approx 2.82$, where $-V_{2,C}$ is the Fisher information about λ for the pure counting experiment (i.e., in the case $s(x) = b(x)$).

Effective z Errors of the Asymptotic Distributions

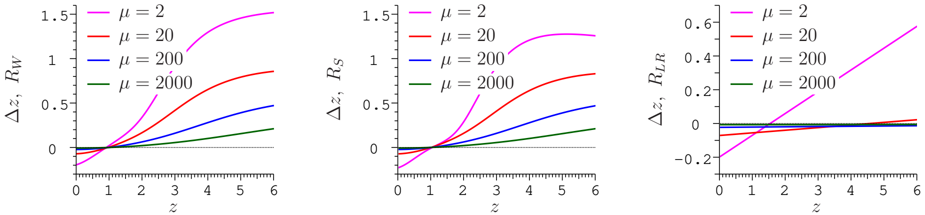
- Transformation
$$z' = \Phi^{-1}(Pr(R < z)), \tag{20}$$
normalizes the distribution of a statistic R. $\Phi(\cdot)$ is the CDF of $\mathcal{N}(0,1)$ and $\Phi^{-1}(\cdot)$ is the quantile function.
- Effective z error of the asymptotic approximation for a statistic with $\mathcal{N}(0,1)$ limiting distribution:
$$\Delta z(z) = z - z'. \tag{21}$$
- For $\mathcal{N}(0,1)$, $\Delta z(z) = 0$.
- To first order, if $Pr(R < z) = \Phi(z) - \frac{d\Phi(z)}{dz} h(z)$, Δz is just $h(z)$.
- In practice, to ensure numerical stability at large z values, we have to calculate instead
$$\Delta z = z - S^{-1}(1 - Pr(R < z)), \tag{22}$$
where $S(z) := 1 - \Phi(z) = \frac{1}{2}\text{erfc}\left(\frac{z}{\sqrt{2}}\right)$ is the survival function of $\mathcal{N}(0,1)$. Of course, subtractive cancellation should also be avoided in the evaluation of $1 - Pr(R < z)$ (i.e., the survival function of R).

Example z Error Predictions to $\mathcal{O}(N^{-3/2})$, Mixture Model



Note different vertical scales in these plots.

Example z Error Predictions to $\mathcal{O}(\mu^{-3/2})$, Poisson Model



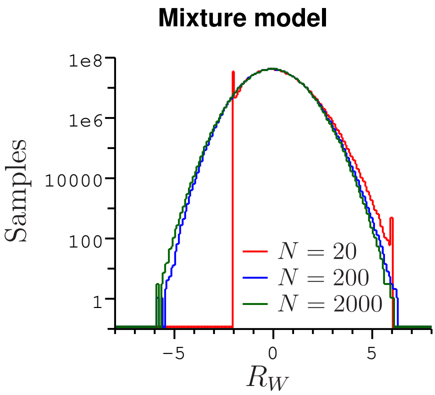
The correspondence between Δz values and p-values at $z = 5$ is as follows (using $p_5 = 1 - \Phi(5) \approx 2.87 \times 10^{-7}$):

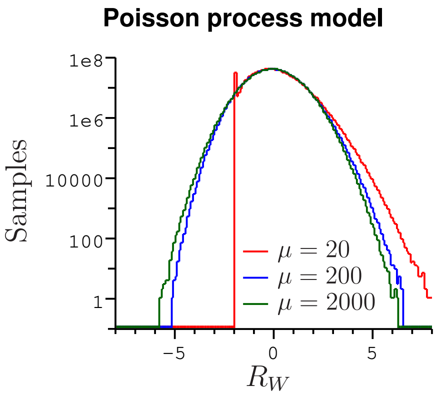
Δz	p-value$/p_5$
-0.01	0.95
0.01	1.05
0.05	1.29
0.1	1.67
0.2	2.77
0.5	11.9
1.0	110

Simulation

- Want to have a reasonable precision of the p-value determination in the $z \sim 5$ region, a traditional threshold for signal discovery claims in high energy physics. Need $\sim 10^9$ samples.
- $M = 10^9$ pseudo-experiments were generated for a number of different N (mixture model) and μ (Poisson process model) settings.
- $\hat{\alpha}$ and $\hat{\lambda}$ were found by maximizing the log-likelihood numerically, and distributions of various statistics were compared with predictions.

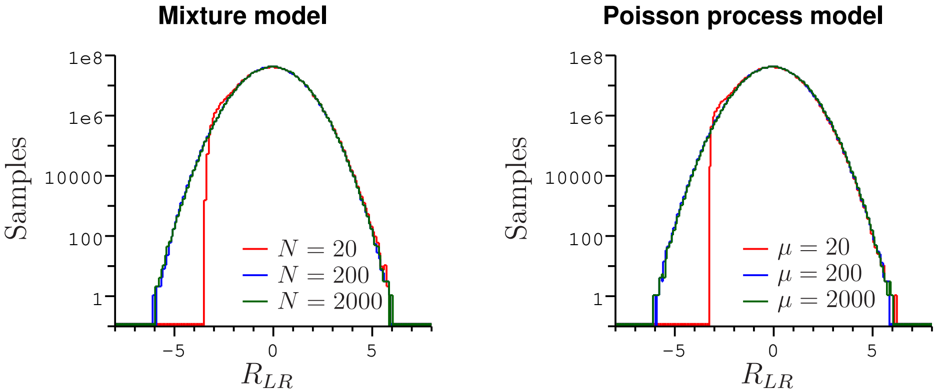
Simulated Distributions of R_W



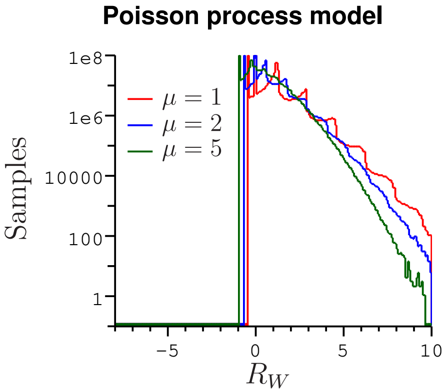


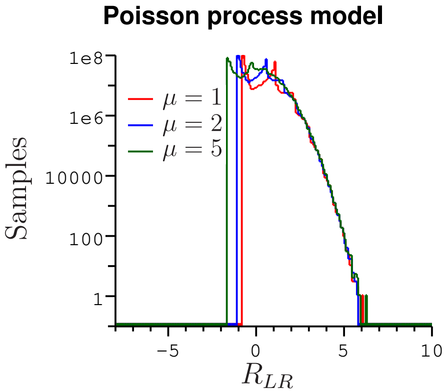
There are "hard" lower and upper limits on $\hat{\alpha}$ ($\hat{\alpha} \in [-0.3345, 1.0]$) and a "hard" lower limit on $\hat{\lambda}$ ($\hat{\lambda} > -0.2507\mu$) stemming from the requirement that the combined signal+background probability density must remain non-negative. The distribution cutoffs, visible for $N = 20$ (left) and $\mu = 20$ (right), are due to these limits.

Simulated Distributions of R_{LR}

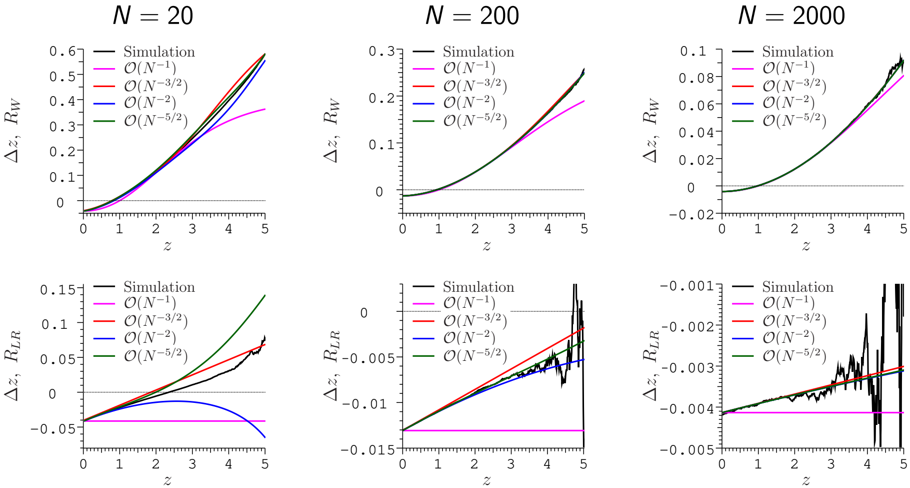


Simulated Distributions for Small μ

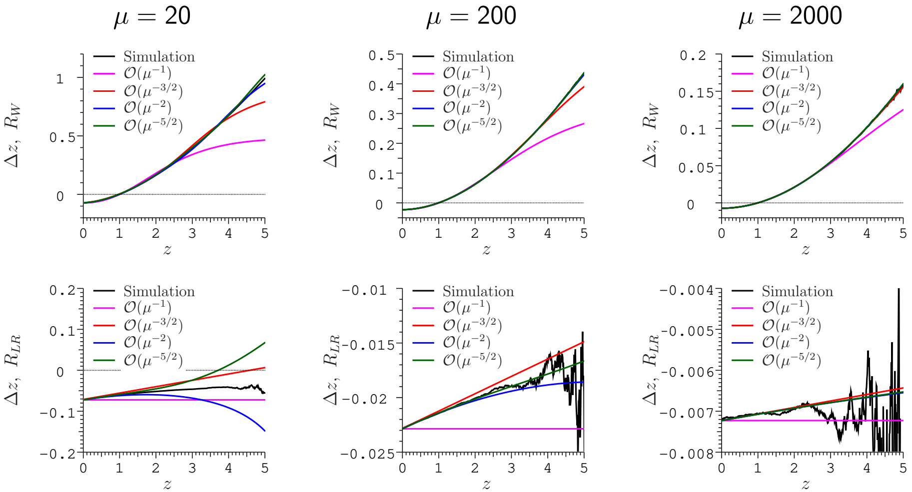




z Errors, Mixture Model

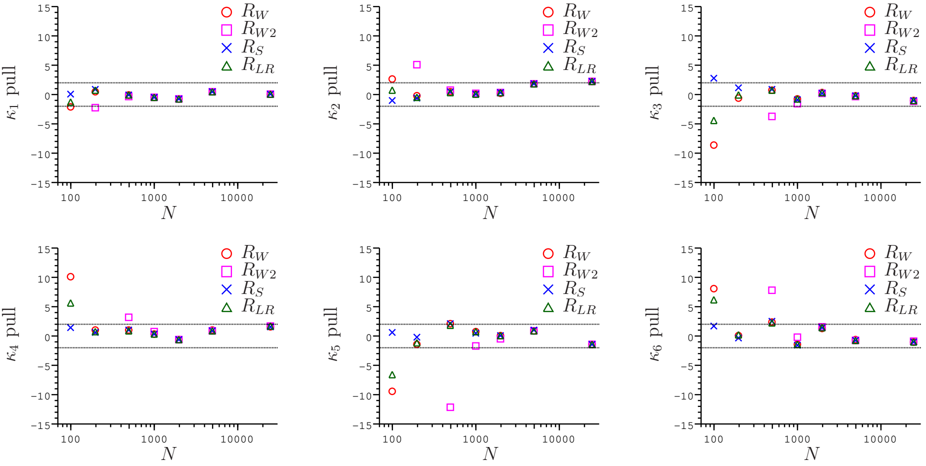


z Errors, Poisson Process Model



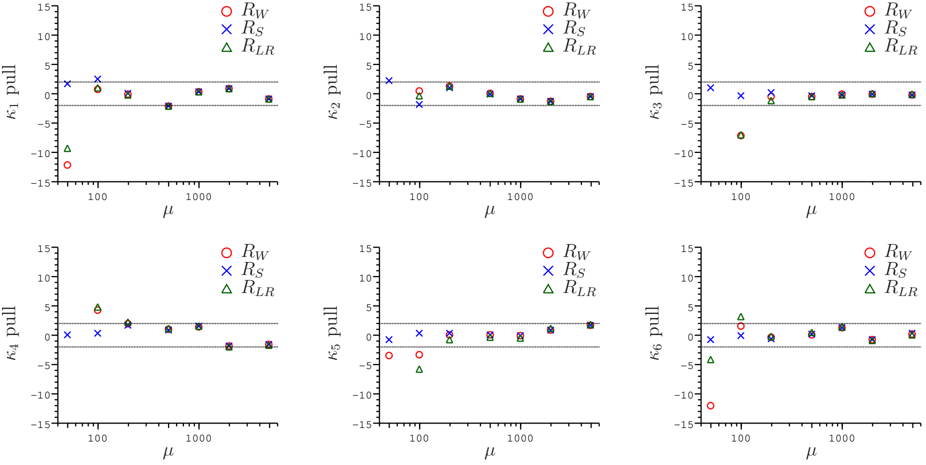
Cumulant Matching to $\mathcal{O}(N^{-5/2})$, Mixture Model

The cumulant pulls are defined as $\frac{k_{i,simulated} - \kappa_{i,predicted}}{\sigma(\kappa_i)_{predicted}}$. Diagnostics based on cumulant matching become unreliable for $N \lessapprox 120$ due to the simulation pile-up at the "hard" lower limit on $\hat{\alpha}$ (i.e., when $\alpha_{true} - \alpha_{min} \lessapprox 5\sigma_{CR}$).

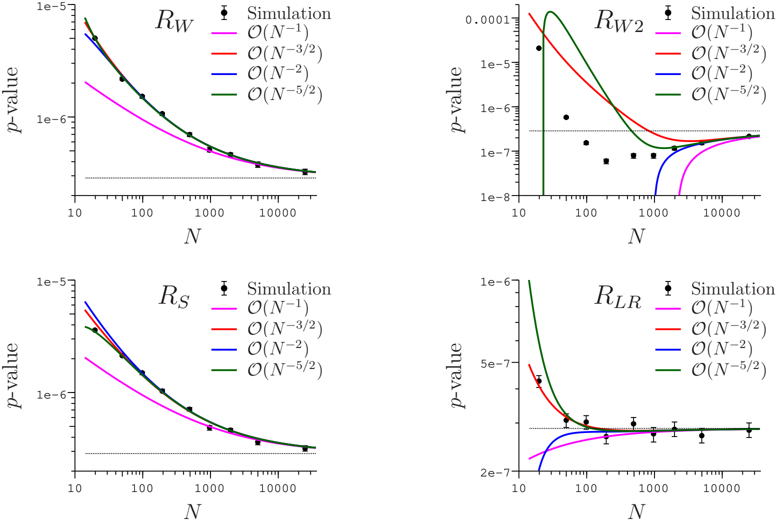


Cumulant Matching to $\mathcal{O}(\mu^{-5/2})$, Poisson Process Model

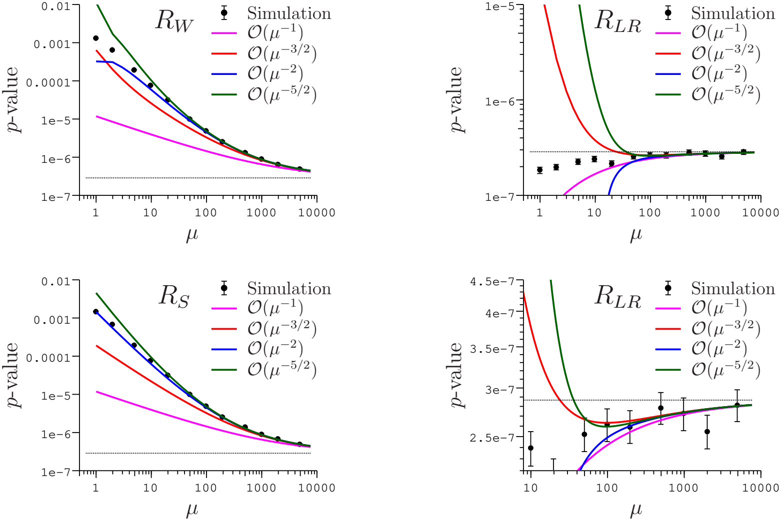
Diagnostics based on cumulant matching become unreliable for $\mu \lessapprox 50$ due to the simulation pile-up at the "hard" lower limit on $\hat{\lambda}$ (i.e., when $\lambda_{true} - \lambda_{min} \lessapprox 5\sigma_{CR}$).



p-values at $z = 5$, Mixture Model



p-values at $z = 5$, Poisson Process Model



Observations

- For the example considered, fourth order Edgeworth asymptotic expansions become highly accurate for modeling R_W, R_S, and R_{LR} distributions when the sample size reaches ~ 100. R_{W2} and R_{W3} need ~ 1000 points.
- Due to its unique cumulant decay rate, R_{LR} exhibits much closer adherence to normality than all other signal testing statistics considered.
- For the mixture model, $\mathcal{O}(N^{-3/2})$ calculation provides excellent p-value approximation for R_{LR} at all sample sizes practically relevant for signal detection at $z \sim 5$. Note that, to $\mathcal{O}(N^{-3/2})$, the R_{LR} distribution is still just $\mathcal{N}(\kappa_1, \kappa_2)$.
- For the Poisson process model, $\mathcal{O}(\mu^{-1/2})$ approximation of R_{LR} (i.e., the leading order) is conservative and works remarkably well down to very small sample sizes.

Note that these observations are valid only for the particular set of shape parameters realized in this example.

Software

- Cumulant calculations of the signal testing statistics up to fourth order in $N^{-1/2}$ (or $\mu^{-1/2}$) are implemented in the NPStat software package, together with the corresponding Edgeworth asymptotic distributions. Relevant functions and classes are mixtureModelCumulants, poissonProcessCumulants, and EdgeworthSeries1D. These calculations take likelihood shape parameters γ, ρ, ξ, and ζ (Eq. 17) as inputs.
- Choice of the optimal method for calculating the shape parameters (according to Eqs. 43 - 51) will, in general, depend on the type (parametric or nonparametric) and shape of the $s(\boldsymbol{x})$ and $b(\boldsymbol{x})$ densities. For example, for 1-d Gaussian $s(x)$ and smooth, parametric $b(x)$, Gauss-Hermite quadrature will be the most appropriate. We did not attempt to devise a generic procedure for calculating the shape parameters, but Gauss-Legendre, Gauss-Hermite, and Fejér quadratures are included in the NPStat package and can be used as building blocks for your own calculations.

Directions to Explore

- Different $s(\boldsymbol{x})$ and $b(\boldsymbol{x})$.
- More simulations in the $z \sim 5$ region.
- Very high order Edgeworth expansions for R_S.
- For the Poisson process model, one can construct expansions conditional on N and then combine them numerically.
- Models not based on the Edgeworth expansions.
- Models not based on matching the cumulants for the complete distribution (e.g., match them only for $z > 0$).
- Incorporation of nuisance parameters into high-order calculations.
- New signal testing staistics (e.g., combine predictability of R_S with the good cumulant control of R_{LR}).

Summary

- For small and moderate sample sizes, deviations from normality for likelihood-based statistics used in signal searches can be significant. Moreover, for narrow signals these deviations are unbounded.
- In comparison with other statistics, deviations from normality are substantially milder for the signed/truncated likelihood ratio statistics. Tests based on likelihood ratio should therefore be preferred.
- It is a good idea to keep these deviations under control. Their influence can be estimated by comparing higher order approximations with the asymptotic formulae for various nuisance parameter values.
- For the LEE determination by the Gross-Vitells method, the random fields can be normalized using Eq. 20. In addition, local significance of the signal test statistic can be adjusted conservatively, leading to a subsequent conservative estimate of the global p-value.
- While the Type 2 error performance of various signal testing statistics was not covered in this talk, practical choice of the statistic should take it into account.

Backup Slides

Higher-Order Asymptotics

- The statements about asymptotic $\mathcal{N}(0,1)$ behavior of various signal testing statistics are valid to $\mathcal{O}(N^{-1/2})$. Informally this means that, for finite N, a corrected quantity $(1+\mathcal{O}(N^{-1/2}))\,R+\mathcal{O}(N^{-1/2})$ can be found which is distributed as stated.
- In many situations, it is possible to construct higher-order approximations explicitly. These approximations, arranged in the powers of $N^{-1/2}$, give us control over the differences between the finite N distribution of a statistic and its limiting behavior.
- The proper small parameter for the Poisson process model is $\mu^{-1/2}$.

The Roadmap for the Mixture Model

1. Represent the log-likelihood derivatives by

$$\left.\frac{d^k \ell(\alpha)}{d\alpha^k}\right|_{\alpha=0} = NV_k + \sqrt{N} Z_k, \tag{23}$$

   where $V_k := \mathrm{E}\left[\frac{d^k \ln p(x|\alpha)}{d\alpha^k}\right]\Big|_{\alpha=0}$ and Z_k is an $\mathcal{O}(1)$ random variable with zero mean. Note that V_k are just model-dependent constants (also $V_1 = 0$ for any model).
2. Construct a sufficiently high order Taylor expansion for $\ell(\alpha)$ at $\alpha = 0$ and solve for $\hat{\alpha}$ in terms of V_k and Z_k. The result will look like

$$\hat{\alpha} = \sum_{m=0}^{M} p_m(Z_1, Z_2, ...) N^{-m/2}, \tag{24}$$

   where $p_m(Z_1, Z_2, ...)$ are certain multivariate polynomials. These polynomials are also complicated functions of V_k but we need to derive them only once. They do not depend on the model or statistic.
3. Using $\hat{\alpha}$ from the previous step, construct the expression for the test statistic of interest in terms of V_k and Z_k.

The Roadmap for the Mixture Model (cont'd)

4. Derive the cumulants of the test statistic of interest in terms of V_k and the joint cumulants of Z_k.
5. Evaluate the joint cumulants of Z_k under the null hypothesis (i.e., $\alpha = 0$). Propagate the results to the cumulants of the test statistic of interest.
6. Using the cumulants of the test statistic, model its distribution with the Edgeworth expansion.

The whole procedure is reminiscent of the quantum mechanics perturbation theory before the Feynman diagrams were invented.

A similar ansatz can be made for the Poisson process model. In this case, various expectations have to be taken w.r.t. both the Poisson distribution of N with parameter μ and the joint null hypothesis distribution of $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_N\}$ conditional on N.

Taylor Series for the Mixture Model Log-Likelihood

Near $\alpha = 0$,

$$\ell(\alpha) = \ell(0) + c_1\alpha + \frac{1}{2!}c_2\alpha^2 + \frac{1}{3!}c_3\alpha^3 + ... \tag{25}$$

$$\ell(0) = \sum_{i=1}^{N} \ln b(x_i), \tag{26}$$

$$c_1 = \left.\frac{d\ell(\alpha)}{d\alpha}\right|_{\alpha=0} = \sum_{i=1}^{N} \frac{s(x_i) - b(x_i)}{b(x_i)}, \tag{27}$$

$$c_2 = \left.\frac{d^2\ell(\alpha)}{d\alpha^2}\right|_{\alpha=0} = -\sum_{i=1}^{N} \frac{(s(x_i) - b(x_i))^2}{b^2(x_i)}, \tag{28}$$

$$c_3 = \left.\frac{d^3\ell(\alpha)}{d\alpha^3}\right|_{\alpha=0} = 2\sum_{i=1}^{N} \frac{(s(x_i) - b(x_i))^3}{b^3(x_i)}, \tag{29}$$

$$c_k = \left.\frac{d^k\ell(\alpha)}{d\alpha^k}\right|_{\alpha=0} = (-1)^{k-1}(k-1)!\sum_{i=1}^{N} \frac{(s(x_i) - b(x_i))^k}{b^k(x_i)}. \tag{30}$$

Asymptotic Expansion for $\hat{\alpha}$ (Model-Independent)

Assume that $c_i = NV_i + \sqrt{N} Z_i$ (as in Eq. 23) and Z_i is a random variable with $E[Z_i] = 0$ and $\mathrm{Var}[Z_i] \sim 1$ (*i.e.*, independent from N). Then the condition $\left.\frac{d\ell(\alpha)}{d\alpha}\right|_{\alpha=\hat{\alpha}} = 0$ leads to

$$c_1 + c_2\hat{\alpha} + \frac{1}{2!} c_3 \hat{\alpha}^2 + \frac{1}{3!} c_4 \hat{\alpha}^3 + \mathcal{O}(N^{-1}) = 0. \tag{31}$$

This equation can be solved for $\hat{\alpha}$ approximately, setting in turn coefficients for each power of N to 0. Taking into account the condition $V_1 = 0$, one obtains (in agreement with Eq. 5.7 in Severini)

$$\begin{aligned}
\sqrt{N}\hat{\alpha} = &- \frac{Z_1}{V_2} \\
&+ N^{-1/2} \left(\frac{Z_1 Z_2}{V_2^2} - \frac{1}{2} \frac{V_3}{V_2^3} Z_1^2 \right) \\
&+ N^{-1} \left[\frac{3}{2} \frac{V_3}{V_2^4} Z_1^2 Z_2 - \frac{1}{2} \frac{Z_1^2 Z_3}{V_2^3} - \frac{Z_1 Z_2^2}{V_2^3} + Z_1^3 \left(\frac{1}{6} \frac{V_4}{V_2^4} - \frac{1}{2} \frac{V_3^2}{V_2^5} \right) \right] \\
&+ \mathcal{O}(N^{-3/2})
\end{aligned} \tag{32}$$

Expansion for $\hat{\alpha}$ to $\mathcal{O}(N^{-5/2})$

Taking couple more steps in the same direction, we arrive at

$$
\begin{aligned}
\sqrt{N}\hat{\alpha} = &- \frac{Z_1}{V_2} \\
&+ N^{-1/2} \left(\frac{Z_1 Z_2}{V_2^2} - \frac{1}{2} \frac{V_3}{V_2^3} Z_1^2 \right) \\
&+ N^{-1} \left[\frac{3}{2} \frac{V_3}{V_2^4} Z_1^2 Z_2 - \frac{1}{2} \frac{Z_1^2 Z_3}{V_2^3} - \frac{Z_1 Z_2^2}{V_2^3} + Z_1^3 \left(\frac{1}{6} \frac{V_4}{V_2^4} - \frac{1}{2} \frac{V_3^2}{V_2^5} \right) \right] \\
&+ N^{-3/2} \left(\begin{array}{l} - \frac{5}{8} \frac{V_3^3 Z_1^4}{V_2^7} + \frac{5}{2} \frac{V_3^2 Z_1^3 Z_2}{V_2^6} + \frac{5}{12} \frac{V_3 V_4 Z_1^4}{V_2^6} - \frac{V_3 Z_1^3 Z_3}{V_2^5} - 3 \frac{V_3 Z_1^2 Z_2^2}{V_2^5} \\ - \frac{2}{3} \frac{V_4 Z_1^3 Z_2}{V_2^5} - \frac{1}{24} \frac{V_5 Z_1^4}{V_2^5} + \frac{1}{6} \frac{Z_1^3 Z_4}{V_2^4} + \frac{3}{2} \frac{Z_1^2 Z_2 Z_3}{V_2^4} + \frac{Z_1 Z_2^3}{V_2^4} \end{array} \right) \\
&+ N^{-2} \left(\begin{array}{l} - \frac{7}{8} \frac{V_3^4 Z_1^5}{V_2^9} + \frac{35}{8} \frac{V_3^3 Z_1^4 Z_2}{V_2^8} + \frac{7}{8} \frac{V_3^2 V_4 Z_1^5}{V_2^8} - \frac{15}{8} \frac{V_3^2 Z_1^4 Z_3}{V_2^7} - \frac{15}{2} \frac{V_3^2 Z_1^3 Z_2^2}{V_2^7} - \frac{5}{2} \frac{V_3 V_4 Z_1^4 Z_2}{V_2^7} - \frac{1}{8} \frac{V_3 V_5 Z_1^5}{V_2^7} \\ - \frac{1}{12} \frac{V_4^2 Z_1^5}{V_2^7} + \frac{5}{12} \frac{V_3 Z_1^4 Z_4}{V_2^6} + 5 \frac{V_3 Z_1^3 Z_2 Z_3}{V_2^6} + 5 \frac{V_3 Z_1^2 Z_2^3}{V_2^6} + \frac{5}{12} \frac{V_4 Z_1^4 Z_3}{V_2^6} + \frac{5}{3} \frac{V_4 Z_1^3 Z_2^2}{V_2^6} + \frac{5}{24} \frac{V_5 Z_1^4 Z_2}{V_2^6} \\ + \frac{1}{120} \frac{V_6 Z_1^5}{V_2^6} - \frac{1}{24} \frac{Z_1^4 Z_5}{V_2^5} - \frac{2}{3} \frac{Z_1^3 Z_2 Z_4}{V_2^5} - \frac{1}{2} \frac{Z_1^3 Z_3^2}{V_2^5} - 3 \frac{Z_1^2 Z_2^2 Z_3}{V_2^5} - \frac{Z_1 Z_2^4}{V_2^5} \end{array} \right) \\
&+ \mathcal{O}(N^{-5/2})
\end{aligned} \tag{33}
$$

There are some hints of structure in this, but it is far from obvious how to make arbitrary order calculations automatic in, let say, a C++ program

Expansion for T_{LR}

$$
\begin{aligned}
T_{LR} = -\frac{Z_1^2}{V_2} &+ N^{-1/2}\left(\frac{Z_1^2 Z_2}{V_2^2} - \frac{1}{3}\frac{V_3 Z_1^3}{V_2^3}\right) \\
&+ N^{-1}\left(\frac{V_3 Z_1^3 Z_2}{V_2^4} - \frac{1}{4}\frac{V_3^2 Z_1^4}{V_2^5} - \frac{1}{3}\frac{Z_1^3 Z_3}{V_2^3} + \frac{1}{12}\frac{V_4 Z_1^4}{V_2^4} - \frac{Z_1^2 Z_2^2}{V_2^3}\right) \\
&+ N^{-3/2}\begin{pmatrix} -\frac{1}{4}\frac{V_3^3 Z_1^5}{V_2^7} + \frac{5}{4}\frac{V_3^2 Z_1^4 Z_2}{V_2^6} + \frac{1}{6}\frac{V_3 V_4 Z_1^5}{V_2^6} - \frac{1}{2}\frac{V_3 Z_1^4 Z_3}{V_2^5} - 2\frac{V_3 Z_1^3 Z_2^2}{V_2^5} \\ -\frac{1}{3}\frac{V_4 Z_1^4 Z_2}{V_2^5} - \frac{1}{60}\frac{V_5 Z_1^5}{V_2^5} + \frac{1}{12}\frac{Z_1^4 Z_4}{V_2^4} + \frac{Z_1^3 Z_2 Z_3}{V_2^4} + \frac{Z_1^2 Z_2^3}{V_2^4} \end{pmatrix} \\
&+ N^{-2}\begin{pmatrix} -\frac{7}{24}\frac{V_3^4 Z_1^6}{V_2^9} + \frac{7}{4}\frac{V_3^3 Z_1^5 Z_2}{V_2^8} + \frac{7}{24}\frac{V_3^2 V_4 Z_1^6}{V_2^8} - \frac{3}{4}\frac{V_3^2 Z_1^5 Z_3}{V_2^7} - \frac{15}{4}\frac{V_3^2 Z_1^4 Z_2^2}{V_2^7} \\ -\frac{V_3 V_4 Z_1^5 Z_2}{V_2^7} - \frac{1}{24}\frac{V_3 V_5 Z_1^6}{V_2^7} - \frac{1}{36}\frac{V_4^2 Z_1^6}{V_2^7} + \frac{1}{6}\frac{V_3 Z_1^5 Z_4}{V_2^6} + \frac{5}{2}\frac{V_3 Z_1^4 Z_2 Z_3}{V_2^6} \\ +\frac{10}{3}\frac{V_3 Z_1^3 Z_2^3}{V_2^6} + \frac{1}{6}\frac{V_4 Z_1^5 Z_3}{V_2^6} + \frac{5}{6}\frac{V_4 Z_1^4 Z_2^2}{V_2^6} + \frac{1}{12}\frac{V_5 Z_1^5 Z_2}{V_2^6} + \frac{1}{360}\frac{V_6 Z_1^6}{V_2^6} \\ -\frac{1}{60}\frac{Z_1^5 Z_5}{V_2^5} - \frac{1}{3}\frac{Z_1^4 Z_2 Z_4}{V_2^5} - \frac{1}{4}\frac{Z_1^4 Z_3^2}{V_2^5} - 2\frac{Z_1^3 Z_2^2 Z_3}{V_2^5} - \frac{Z_1^2 Z_2^4}{V_2^5} \end{pmatrix} \\
&+ \mathcal{O}(N^{-5/2})
\end{aligned} \tag{34}
$$

The leading terms of this expansion are consistent with Severini (in his book, it is derived to $\mathcal{O}(N^{-3/2})$)

A Comment on the Higher-Order Distribution of T_{LR}

- For positive values of T_{TLR} (truncated T_{LR}), $Pr(T_{TLR} < z^2) = Pr(R_{LR} < z)$, where $R_{LR} = \sqrt{T_{TLR}}$.
- The following second order treatment of T_{LR} (not truncated) is especially effective. Construct the adjusted statistic: $Pr(T_{BLR} < z^2) := Pr\left(T_{LR} < z^2 \left(1 + \frac{b}{N}\right)^{-1}\right)$. T_{BLR} is known as the Bartlett-adjusted likelihood ratio statistic, where b is the so-called Bartlett coefficient. In terms of the second-order cumulants of R_{LR}, $b = N(\kappa_1^2 + \kappa_2 - 1)$. Then

$$Pr(T_{BLR} < z^2) = \text{erf}\left(\frac{z}{\sqrt{2}}\right) + O(N^{-2}). \tag{35}$$

That is, T_{BLR} has a χ_1^2 distribution to order $O(N^{-2})$: it turns out that $O(N^{-3/2})$ terms cancel out. However, to use T_{BLR}, you *must* treat negative and positive signal fractions on equal footing (not very practical for signal searches).

Expansion for R_{LR}

$R_{LR} = \frac{Z_1}{\sqrt{-V_2}}(1 + \Delta_{LR})$, where

$$
\begin{aligned}
\Delta_{LR} = N^{-1/2} & \left(\frac{1}{6} \frac{V_3 Z_1}{V_2^2} - \frac{1}{2} \frac{Z_2}{V_2} \right) \\
+ N^{-1} & \left(\frac{1}{9} \frac{V_3^2 Z_1^2}{V_2^4} - \frac{5}{12} \frac{V_3 Z_1 Z_2}{V_2^3} - \frac{1}{24} \frac{V_4 Z_1^2}{V_2^3} + \frac{1}{6} \frac{Z_1 Z_3}{V_2^2} + \frac{3}{8} \frac{Z_2^2}{V_2^2} \right) \\
+ N^{-3/2} & \left(\begin{array}{l} \frac{23}{216} \frac{V_3^3 Z_1^3}{V_2^6} - \frac{1}{2} \frac{V_3^2 Z_1^2 Z_2}{V_2^5} - \frac{11}{144} \frac{V_3 V_4 Z_1^3}{V_2^5} + \frac{2}{9} \frac{V_3 Z_1^2 Z_3}{V_2^4} + \frac{35}{48} \frac{V_3 Z_1 Z_2^2}{V_2^4} \\ + \frac{7}{48} \frac{V_4 Z_1^2 Z_2}{V_2^4} + \frac{1}{120} \frac{V_5 Z_1^3}{V_2^4} - \frac{1}{24} \frac{Z_1^2 Z_4}{V_2^3} - \frac{5}{12} \frac{Z_1 Z_2 Z_3}{V_2^3} - \frac{5}{16} \frac{Z_2^3}{V_2^3} \end{array} \right) \\
+ N^{-2} & \left(\begin{array}{l} \frac{79}{648} \frac{V_3^4 Z_1^4}{V_2^8} - \frac{299}{432} \frac{V_3^3 Z_1^3 Z_2}{V_2^7} - \frac{37}{288} \frac{V_3^2 V_4 Z_1^4}{V_2^7} + \frac{23}{72} \frac{V_3^2 Z_1^3 Z_3}{V_2^6} + \frac{11}{8} \frac{V_3^2 Z_1^2 Z_2^2}{V_2^6} \\ + \frac{121}{288} \frac{V_3 V_4 Z_1^3 Z_2}{V_2^6} + \frac{7}{360} \frac{V_3 V_5 Z_1^4}{V_2^6} + \frac{5}{384} \frac{V_4^2 Z_1^4}{V_2^6} - \frac{11}{144} \frac{V_3 Z_1^3 Z_4}{V_2^5} - \frac{V_3 Z_1^2 Z_2 Z_3}{V_2^5} \\ - \frac{35}{32} \frac{V_3 Z_1 Z_2^3}{V_2^5} - \frac{11}{144} \frac{V_4 Z_1^3 Z_3}{V_2^5} - \frac{21}{64} \frac{V_4 Z_1^2 Z_2^2}{V_2^5} - \frac{3}{80} \frac{V_5 Z_1^3 Z_2}{V_2^5} - \frac{1}{720} \frac{V_6 Z_1^4}{V_2^5} \\ + \frac{1}{120} \frac{Z_1^3 Z_5}{V_2^4} + \frac{7}{48} \frac{Z_1^2 Z_2 Z_4}{V_2^4} + \frac{1}{9} \frac{Z_1^2 Z_3^2}{V_2^4} + \frac{35}{48} \frac{Z_1 Z_2^2 Z_3}{V_2^4} + \frac{35}{128} \frac{Z_2^4}{V_2^4} \end{array} \right) \\
+ & \mathcal{O}(N^{-5/2})
\end{aligned}
\tag{36}
$$

It appears that this expansion is not correct in the Severini's book.

Expansion for R_{W2}

$R_{W2} = \frac{Z_1}{\sqrt{-V_2}}(1 + \Delta_{W2})$, where

$$\begin{aligned}
\Delta_{W2} = {} & N^{-1/2}\left(-\frac{1}{2}\frac{Z_2}{V_2}\right) \\
& + N^{-1}\left(-\frac{1}{8}\frac{V_3^2 Z_1^2}{V_2^4} + \frac{1}{12}\frac{V_4 Z_1^2}{V_2^3} + \frac{3}{8}\frac{Z_2^2}{V_2^2}\right) \\
& + N^{-3/2}\left(\begin{array}{l} -\frac{1}{4}\frac{V_3^3 Z_1^3}{V_2^6} + \frac{9}{16}\frac{V_3^2 Z_1^2 Z_2}{V_2^5} + \frac{1}{4}\frac{V_3 V_4 Z_1^3}{V_2^5} - \frac{1}{4}\frac{V_3 Z_1^2 Z_3}{V_2^4} \\ -\frac{7}{24}\frac{V_4 Z_1^2 Z_2}{V_2^4} - \frac{1}{24}\frac{V_5 Z_1^3}{V_2^4} + \frac{1}{12}\frac{Z_1^2 Z_4}{V_2^3} - \frac{5}{16}\frac{Z_2^3}{V_2^3} \end{array}\right) \\
& + N^{-2}\left(\begin{array}{l} -\frac{57}{128}\frac{V_3^4 Z_1^4}{V_2^8} + \frac{13}{8}\frac{V_3^3 Z_1^3 Z_2}{V_2^7} + \frac{19}{32}\frac{V_3^2 V_4 Z_1^4}{V_2^7} - \frac{3}{4}\frac{V_3^2 Z_1^3 Z_3}{V_2^6} - \frac{99}{64}\frac{V_3^2 Z_1^2 Z_2^2}{V_2^6} - \frac{11}{8}\frac{V_3 V_4 Z_1^3 Z_2}{V_2^6} \\ -\frac{1}{8}\frac{V_3 V_5 Z_1^4}{V_2^6} - \frac{7}{96}\frac{V_4^2 Z_1^4}{V_2^6} + \frac{1}{4}\frac{V_3 Z_1^3 Z_4}{V_2^5} + \frac{9}{8}\frac{V_3 Z_1^2 Z_2 Z_3}{V_2^5} + \frac{1}{4}\frac{V_4 Z_1^3 Z_3}{V_2^5} + \frac{21}{32}\frac{V_4 Z_1^2 Z_2^2}{V_2^5} \\ +\frac{3}{16}\frac{V_5 Z_1^3 Z_2}{V_2^5} + \frac{1}{80}\frac{V_6 Z_1^4}{V_2^5} - \frac{1}{24}\frac{Z_1^3 Z_5}{V_2^4} - \frac{7}{24}\frac{Z_1^2 Z_2 Z_4}{V_2^4} - \frac{1}{8}\frac{Z_1^2 Z_3^2}{V_2^4} + \frac{35}{128}\frac{Z_2^4}{V_2^4} \end{array}\right) \\
& + \mathcal{O}(N^{-5/2})
\end{aligned} \tag{37}$$

Expansion for R_{W3}

$$
\begin{aligned}
\Delta_{W3} ={}& N^{-1/2}\left(-\frac{1}{2}\,\frac{Z_2}{V_2}\right) + N^{-1}\left(-\frac{1}{8}\,\frac{V_3^2 Z_1^2}{V_2^4} + \frac{5}{72}\,\frac{V_3^2}{V_2^3} + \frac{1}{12}\,\frac{V_4 Z_1^2}{V_2^3} - \frac{1}{24}\,\frac{V_4}{V_2^2} + \frac{3}{8}\,\frac{Z_2^2}{V_2^2}\right) \\
&+ N^{-3/2}\left(\begin{array}{l}
-\frac{1}{4}\,\frac{V_3^3 Z_1^3}{V_2^6} + \frac{5}{24}\,\frac{V_3^3 Z_1}{V_2^5} + \frac{9}{16}\,\frac{V_3^2 Z_1^2 Z_2}{V_2^5} + \frac{1}{4}\,\frac{V_3 V_4 Z_1^3}{V_2^5} - \frac{35}{144}\,\frac{V_3^2 Z_2}{V_2^4} - \frac{2}{9}\,\frac{V_3 V_4 Z_1}{V_2^4} - \frac{1}{4}\,\frac{V_3 Z_1^2 Z_3}{V_2^4} \\
-\frac{7}{24}\,\frac{V_4 Z_1^2 Z_2}{V_2^4} - \frac{1}{24}\,\frac{V_5 Z_1^3}{V_2^4} + \frac{5}{36}\,\frac{V_3 Z_3}{V_2^3} + \frac{5}{48}\,\frac{V_4 Z_2}{V_2^3} + \frac{1}{24}\,\frac{V_5 Z_1}{V_2^3} + \frac{1}{12}\,\frac{Z_1^2 Z_4}{V_2^3} - \frac{5}{16}\,\frac{Z_2^3}{V_2^3} - \frac{1}{24}\,\frac{Z_4}{V_2^2}
\end{array}\right) \\
&+ N^{-2}\left(\begin{array}{l}
-\frac{57}{128}\,\frac{V_3^4 Z_1^4}{V_2^8} + \frac{295}{576}\,\frac{V_3^4 Z_1^2}{V_2^7} + \frac{13}{8}\,\frac{V_3^3 Z_1^3 Z_2}{V_2^7} + \frac{19}{32}\,\frac{V_3^2 V_4 Z_1^4}{V_2^7} - \frac{181}{10368}\,\frac{V_3^4}{V_2^6} - \frac{55}{48}\,\frac{V_3^3 Z_1 Z_2}{V_2^6} - \frac{7}{96}\,\frac{V_4^2 Z_1^4}{V_2^6} \\
-\frac{1289}{1728}\,\frac{V_3^2 V_4 Z_1^2}{V_2^6} - \frac{3}{4}\,\frac{V_3^2 Z_1^3 Z_3}{V_2^6} - \frac{99}{64}\,\frac{V_3^2 Z_1^2 Z_2^2}{V_2^6} - \frac{11}{8}\,\frac{V_3 V_4 Z_1^3 Z_2}{V_2^6} - \frac{1}{8}\,\frac{V_3 V_5 Z_1^4}{V_2^6} + \frac{53}{1728}\,\frac{V_3^2 V_4}{V_2^5} \\
+\frac{5}{8}\,\frac{V_3^2 Z_1 Z_3}{V_2^5} + \frac{35}{64}\,\frac{V_3^2 Z_2^2}{V_2^5} + \frac{V_3 V_4 Z_1 Z_2}{V_2^5} + \frac{25}{144}\,\frac{V_3 V_5 Z_1^2}{V_2^5} + \frac{1}{4}\,\frac{V_3 Z_1^3 Z_4}{V_2^5} + \frac{9}{8}\,\frac{V_3 Z_1^2 Z_2 Z_3}{V_2^5} + \frac{31}{288}\,\frac{V_4^2 Z_1^2}{V_2^5} \\
+\frac{1}{4}\,\frac{V_4 Z_1^3 Z_3}{V_2^5} + \frac{21}{32}\,\frac{V_4 Z_1^2 Z_2^2}{V_2^5} + \frac{3}{16}\,\frac{V_5 Z_1^3 Z_2}{V_2^5} + \frac{1}{80}\,\frac{V_6 Z_1^4}{V_2^5} - \frac{7}{720}\,\frac{V_3 V_5}{V_2^4} - \frac{2}{9}\,\frac{V_3 Z_1 Z_4}{V_2^4} - \frac{35}{72}\,\frac{V_3 Z_2 Z_3}{V_2^4} \\
-\frac{5}{1152}\,\frac{V_4^2}{V_2^4} - \frac{2}{9}\,\frac{V_4 Z_1 Z_3}{V_2^4} - \frac{35}{192}\,\frac{V_4 Z_2^2}{V_2^4} - \frac{7}{48}\,\frac{V_5 Z_1 Z_2}{V_2^4} - \frac{1}{48}\,\frac{V_6 Z_1^2}{V_2^4} - \frac{1}{24}\,\frac{Z_1^3 Z_5}{V_2^4} - \frac{7}{24}\,\frac{Z_1^2 Z_2 Z_4}{V_2^4} \\
-\frac{1}{8}\,\frac{Z_1^2 Z_3^2}{V_2^4} + \frac{35}{128}\,\frac{Z_2^4}{V_2^4} + \frac{1}{720}\,\frac{V_6}{V_2^3} + \frac{1}{24}\,\frac{Z_1 Z_5}{V_2^3} + \frac{5}{48}\,\frac{Z_2 Z_4}{V_2^3} + \frac{5}{72}\,\frac{Z_3^2}{V_2^3}
\end{array}\right) \\
&+ \mathcal{O}(N^{-5/2})
\end{aligned}
\tag{38}
$$

Asymptotic Expansions for Other Statistics

- Divide the expansion for $\hat{\alpha}$ by σ_{CR} ($\sigma_{CR} = 1/\sqrt{-NV_2}$) in order to obtain the expansion for R_W.
- $R_S = \frac{Z_1}{\sqrt{-V_2}}$ *to all orders.*

Cumulants

- A good introduction into cumulant theory can be found in Wikipedia.
- Cumulants are combinations of (univariate or multivariate) distribution moments which are added when two independent random variables with the same distribution are added.
- For univariate distributions, there is a unique mapping between first m distribution moments and first m cumulants. For multivariate distributions, the mapping involves all joint cumulants and moments up to the order given.
- Relationships between cumulants and moments can be determined automatically (from the Taylor expansion of the cumulant generating function or from Bell polynomials) by a symbolic algebra system.
- For $\mathcal{N}(0,1)$, $\kappa_2 = 1$ while all other cumulants are 0.

Edgeworth Expansion for Typical R to $\mathcal{O}(N^{-5/2})$

Notation: $\phi(z)$ is the density of $\mathcal{N}(0,1)$ and $\Phi(z)$ is the CDF of $\mathcal{N}(0,1)$. Expansion w.r.t. $\mathcal{N}(0,1)$ is given by

$$\begin{aligned} Pr(R < z) = \Phi(z) - \phi(z)\Bigg[& \kappa_1 + H_1(z)\left(\frac{\kappa_1^2}{2} + \frac{\kappa_2 - 1}{2}\right) + H_2(z)\left(\frac{\kappa_1^3}{6} + \frac{\kappa_1(\kappa_2 - 1)}{2} + \frac{\kappa_3}{6}\right) \\ & + H_3(z)\left(\frac{\kappa_1^4}{24} + \frac{\kappa_1^2(\kappa_2 - 1)}{4} + \frac{\kappa_1\kappa_3}{6} + \frac{(\kappa_2 - 1)^2}{8} + \frac{\kappa_4}{24}\right) \\ & + H_4(z)\left(\frac{\kappa_1^2\kappa_3}{12} + \frac{\kappa_1\kappa_4}{24} + \frac{(\kappa_2 - 1)\kappa_3}{12} + \frac{\kappa_5}{120}\right) \\ & + H_5(z)\left(\frac{\kappa_1^3\kappa_3}{36} + \frac{\kappa_1^2\kappa_4}{48} + \frac{\kappa_1(\kappa_2 - 1)\kappa_3}{12} + \frac{\kappa_1\kappa_5}{120} + \frac{(\kappa_2 - 1)\kappa_4}{48} + \frac{\kappa_3^2}{72} + \frac{\kappa_6}{720}\right) \\ & + H_6(z)\left(\frac{\kappa_1\kappa_3^2}{72} + \frac{\kappa_3\kappa_4}{144}\right) \\ & + H_7(z)\left(\frac{\kappa_1^2\kappa_3^2}{144} + \frac{\kappa_1\kappa_3\kappa_4}{144} + \frac{(\kappa_2 - 1)\kappa_3^2}{144} + \frac{\kappa_3\kappa_5}{720} + \frac{\kappa_4^2}{1152}\right) \\ & + H_8(z)\frac{\kappa_3^3}{1296} + H_9(z)\left(\frac{\kappa_1\kappa_3^3}{1296} + \frac{\kappa_3^2\kappa_4}{1728}\right) + H_{11}(z)\frac{\kappa_3^4}{31104} + \mathcal{O}(N^{-5/2})\Bigg] \end{aligned} \quad (39)$$

For this distribution, the four leading cumulants are matched exactly, while κ_5 and κ_6 are matched to $\mathcal{O}(N^{-5/2})$. The actual fifth cumulant of this distribution is $\kappa_5 - \kappa_1^5 - 10\kappa_1^3(\kappa_2 - 1) - 15\kappa_1(\kappa_2 - 1)^2$.

Edgeworth Expansion for Typical R w.r.t. $\mathcal{N}(\kappa_1, \kappa_2)$

$$
\begin{aligned}
Pr(R < z) = \Phi\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) - \phi\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) \Bigg[& \frac{\kappa_3}{6\kappa_2^{3/2}} H_2\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) + \frac{\kappa_4}{24\kappa_2^2} H_3\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) \\
& + \frac{\kappa_5}{120\kappa_2^{5/2}} H_4\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) + \frac{\left(10\kappa_3^2 + \kappa_6\right)}{720\kappa_2^3} H_5\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) \\
& + \frac{\kappa_3\kappa_4}{144\kappa_2^{7/2}} H_6\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) + \frac{\left(8\kappa_3\kappa_5 + 5\kappa_4^2\right)}{5760\kappa_2^4} H_7\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) \\
& + \frac{\kappa_3^3}{1296\kappa_2^{9/2}} H_8\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) + \frac{\kappa_3^2\kappa_4}{1728\kappa_2^5} H_9\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) \\
& + \frac{\kappa_3^4}{31104\kappa_2^6} H_{11}\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) + \mathcal{O}(N^{-5/2}) \Bigg]
\end{aligned}
\tag{40}
$$

For this distribution, the six leading cumulants are matched exactly, actual $\kappa_7 = 0$, $\kappa_8 = 0$, $\kappa_9 = -126\kappa_4\kappa_5 - 84\kappa_3\kappa_6$.

Edgeworth Expansions for R_{LR} to $\mathcal{O}(N^{-5/2})$

- Expansion w.r.t. $\mathcal{N}(0,1)$

$$Pr(R_{LR} < z) = \Phi(z) - \phi(z)\Big[\kappa_1 + H_1(z)\left(\frac{\kappa_1^2}{2} + \frac{\kappa_2 - 1}{2}\right) + H_2(z)\left(\frac{\kappa_1^3}{6} + \frac{\kappa_1(\kappa_2 - 1)}{2} + \frac{\kappa_3}{6}\right) + H_3(z)\left(\frac{\kappa_1^4}{24} + \frac{\kappa_1^2(\kappa_2 - 1)}{4} + \frac{\kappa_1\kappa_3}{6} + \frac{(\kappa_2 - 1)^2}{8} + \frac{\kappa_4}{24}\right) + \mathcal{O}(N^{-5/2})\Big] \quad (41)$$

For this distribution, the four leading cumulants are matched exactly. Actual $\kappa_5 = -(\kappa_1^5 + 10\kappa_1^3(\kappa_2 - 1) + 15\kappa_1(\kappa_2 - 1)^2 + 10\kappa_1^2\kappa_3 + 10(\kappa_2 - 1)\kappa_3 + 5\kappa_1\kappa_4)$.

- Expansion w.r.t. $\mathcal{N}(\kappa_1, \kappa_2)$

$$Pr(R_{LR} < z) = \Phi\left(\frac{z - \kappa_1}{\sqrt{\kappa_2}}\right) - \phi\left(\frac{z - \kappa_1}{\sqrt{\kappa_2}}\right)\left[\frac{\kappa_3}{6\kappa_2^{3/2}} H_2\left(\frac{z - \kappa_1}{\sqrt{\kappa_2}}\right) + \frac{\kappa_4}{24\kappa_2^2} H_3\left(\frac{z - \kappa_1}{\sqrt{\kappa_2}}\right) + \mathcal{O}(N^{-5/2})\right] \quad (42)$$

For this distribution, $\kappa_1, ..., \kappa_4$ are matched exactly, actual $\kappa_5 = 0$, $\kappa_6 = -10\kappa_3^3$.

- "Typical R" expansions can also be used for R_{LR} by setting κ_5 and κ_6 formal parameters to 0. This leads to an alternative behavior of actual higher order cumulants.

Shape Parameters for the Mixture Model

E_b denotes "expectation under the null": $\mathrm{E}_b[q] := \int q(x)b(x)dx$.
E_s denotes "expectation under the signal": $\mathrm{E}_s[q] := \int q(x)s(x)dx$.
Note that $\mathrm{E}_b\left[\frac{s^k}{b^k}\right] = \mathrm{E}_s\left[\frac{s^{k-1}}{b^{k-1}}\right]$. For the mixture model, $V_2 = 1 - \mathrm{E}_s\left[\frac{s}{b}\right]$.

$$\gamma = \frac{V_3}{2\,(-V_2)^{3/2}} = \frac{\mathrm{E}_b\left[2\left(\frac{s-b}{b}\right)^3\right]}{2\,(-V_2)^{3/2}} = \frac{\mathrm{E}_s\left[\frac{s^2}{b^2}\right] - 3\mathrm{E}_s\left[\frac{s}{b}\right] + 2}{\left(\mathrm{E}_s\left[\frac{s}{b}\right] - 1\right)^{3/2}} \tag{43}$$

$$\rho = -\frac{V_4}{6V_2^2} = -\frac{\mathrm{E}_b\left[-6\left(\frac{s-b}{b}\right)^4\right]}{6V_2^2} = \frac{\mathrm{E}_s\left[\frac{s^3}{b^3}\right] - 4\mathrm{E}_s\left[\frac{s^2}{b^2}\right] + 6\mathrm{E}_s\left[\frac{s}{b}\right] - 3}{\left(\mathrm{E}_s\left[\frac{s}{b}\right] - 1\right)^2} \tag{44}$$

$$\xi = \frac{V_5}{24\,(-V_2)^{5/2}} = \frac{\mathrm{E}_s\left[\frac{s^4}{b^4}\right] - 5\mathrm{E}_s\left[\frac{s^3}{b^3}\right] + 10\mathrm{E}_s\left[\frac{s^2}{b^2}\right] - 10\mathrm{E}_s\left[\frac{s}{b}\right] + 4}{\left(\mathrm{E}_s\left[\frac{s}{b}\right] - 1\right)^{5/2}} \tag{45}$$

$$\zeta = \frac{V_6}{120\,V_2^3} = \frac{\mathrm{E}_s\left[\frac{s^5}{b^5}\right] - 6\mathrm{E}_s\left[\frac{s^4}{b^4}\right] + 15\mathrm{E}_s\left[\frac{s^3}{b^3}\right] - 20\mathrm{E}_s\left[\frac{s^2}{b^2}\right] + 15\mathrm{E}_s\left[\frac{s}{b}\right] - 5}{\left(\mathrm{E}_s\left[\frac{s}{b}\right] - 1\right)^3} \tag{46}$$

Shape Parameters for the Poisson Process Model

$$V_2 = -\frac{\mathrm{E}_s\left[\frac{s}{b}\right]}{\mu^2} \quad (47)$$

$$\gamma = \frac{V_3}{2\,(-V_2)^{3/2}} = \frac{\mathrm{E}_s\left[\frac{s^2}{b^2}\right]}{\left(\mathrm{E}_s\left[\frac{s}{b}\right]\right)^{3/2}} \quad (48)$$

$$\rho = -\frac{V_4}{6\,V_2^2} = \frac{\mathrm{E}_s\left[\frac{s^3}{b^3}\right]}{\left(\mathrm{E}_s\left[\frac{s}{b}\right]\right)^{2}} \quad (49)$$

$$\xi = \frac{V_5}{24\,(-V_2)^{5/2}} = \frac{\mathrm{E}_s\left[\frac{s^4}{b^4}\right]}{\left(\mathrm{E}_s\left[\frac{s}{b}\right]\right)^{5/2}} \quad (50)$$

$$\zeta = \frac{V_6}{120\,V_2^3} = \frac{\mathrm{E}_s\left[\frac{s^5}{b^5}\right]}{\left(\mathrm{E}_s\left[\frac{s}{b}\right]\right)^{3}} \quad (51)$$

For pure counting experiments, higher-order approximations of various statistics can be obtained by setting $s = b$ and, consequently, $\gamma = \rho = \xi = \zeta = 1$.

Cumulants to $\mathcal{O}(N^{-5/2})$ for the Mixture Model

Signed Wald statistic, R_W:

$$\kappa_1 = \frac{1}{N^{3/2}}(-\gamma^3 + 2\gamma\rho - \xi),$$
$$\kappa_2 = 1 + \frac{1}{N}(\rho - \gamma^2 - 1) + \frac{1}{N^2}(-10\gamma^4 + 30\gamma^2\rho + 3\gamma^2 - 22\gamma\xi - 8\rho^2 - 3\rho + 10\zeta + 1),$$
$$\kappa_3 = \frac{1}{\sqrt{N}}\gamma + \frac{1}{N^{3/2}}(-12\gamma^3 + 21\gamma\rho - 3\gamma - 9\xi),$$
$$\kappa_4 = \frac{1}{N}(\rho - 3) + \frac{1}{N^2}(-114\gamma^4 + 288\gamma^2\rho + 12\gamma^2 - 180\gamma\xi - 60\rho^2 - 18\rho + 66\zeta + 12),$$
$$\kappa_5 = \frac{1}{N^{3/2}}(-30\gamma^3 + 50\gamma\rho - 20\gamma - 19\xi),$$
$$\kappa_6 = \frac{1}{N^2}(-420\gamma^4 + 930\gamma^2\rho - 40\gamma^2 - 510\gamma\xi - 150\rho^2 - 45\rho + 151\zeta + 60).$$

Signed Wald statistic using observed information, R_{W2}:

$$\kappa_1 = \frac{1}{\sqrt{N}}\left(-\frac{\gamma}{2}\right) + \frac{1}{N^{3/2}}\left(\frac{\gamma^3}{4} + \frac{5\gamma\rho}{16} + \frac{3\gamma}{16} - \frac{9\xi}{8}\right),$$
$$\kappa_2 = 1 + \frac{1}{N}\left(3\rho - \frac{5\gamma^2}{4}\right) + \frac{1}{N^2}\left(\frac{41\gamma^4}{4} - \frac{443\gamma^2\rho}{16} + \frac{19\gamma^2}{16} + \frac{55\gamma\xi}{8} + 10\rho^2 - 3\rho + 6\zeta\right),$$
$$\kappa_3 = \frac{1}{\sqrt{N}}(-2\gamma) + \frac{1}{N^{3/2}}\left(\frac{25\gamma^3}{8} + \frac{3\gamma\rho}{4} + \frac{3\gamma}{4} - 15\xi\right),$$
$$\kappa_4 = \frac{1}{N}(10\rho) + \frac{1}{N^2}\left(\frac{243\gamma^4}{8} - 132\gamma^2\rho - \gamma^2 + 45\gamma\xi + 72\rho^2 - 10\rho + 88\zeta\right),$$
$$\kappa_5 = \frac{1}{N^{3/2}}(15\gamma^3 - 40\gamma\rho - 54\xi),$$
$$\kappa_6 = \frac{1}{N^2}(-60\gamma^4 - 150\gamma^2\rho - 10\gamma^2 + 330\gamma\xi + 360\rho^2 + 376\zeta).$$

Cumulants to $\mathcal{O}(N^{-5/2})$ for the Mixture Model (cont'd)

Signed Wald statistic using $\Delta\ell(\alpha) = -1/2$ uncertainty, R_{W3}:

$$\kappa_1 = \frac{1}{\sqrt{N}}\left(-\frac{\gamma}{2}\right) + \frac{1}{N^{3/2}}\left(-\frac{4\gamma^3}{9} + \frac{259\gamma\rho}{144} + \frac{3\gamma}{16} - \frac{15\xi}{8}\right),$$

$$\kappa_2 = 1 + \frac{1}{N}\left(\frac{7\rho}{2} - \frac{65\gamma^2}{36}\right) + \frac{1}{N^2}\left(\frac{170\gamma^4}{27} - \frac{583\gamma^2\rho}{48} + \frac{251\gamma^2}{144} - \frac{259\gamma\xi}{40} + \frac{131\rho^2}{36} - \frac{7\rho}{2} + \frac{124\zeta}{9}\right),$$

$$\kappa_3 = \frac{1}{\sqrt{N}}(-2\gamma) + \frac{1}{N^{3/2}}\left(-\frac{5\gamma^3}{24} + \frac{107\gamma\rho}{12} + \frac{3\gamma}{4} - \frac{39\xi}{2}\right),$$

$$\kappa_4 = \frac{1}{N}(10\rho) + \frac{1}{N^2}\left(\frac{163\gamma^4}{8} - \frac{748\gamma^2\rho}{9} - \gamma^2 - 4\gamma\xi + \frac{130\rho^2}{3} - 10\rho + 124\zeta\right),$$

$$\kappa_5 = \frac{1}{N^{3/2}}(15\gamma^3 - 40\gamma\rho - 54\xi),$$

$$\kappa_6 = \frac{1}{N^2}(-60\gamma^4 - 150\gamma^2\rho - 10\gamma^2 + 330\gamma\xi + 360\rho^2 + 376\zeta).$$

Cumulants to $\mathcal{O}(N^{-5/2})$ for the Mixture Model (cont'd)

Signed score statistic, R_S:

$$\kappa_1 = 0,$$
$$\kappa_2 = 1,$$
$$\kappa_3 = \frac{1}{\sqrt{N}}\gamma,$$
$$\kappa_4 = \frac{1}{N}(\rho - 3),$$
$$\kappa_5 = \frac{1}{N^{3/2}}(\xi - 10\gamma),$$
$$\kappa_6 = \frac{1}{N^2}(30 - 10\gamma^2 - 15\rho + \zeta).$$

Signed likelihood ratio statistic, R_{LR}:

$$\kappa_1 = \frac{1}{\sqrt{N}}\left(-\frac{\gamma}{6}\right) + \frac{1}{N^{3/2}}\left(-\frac{\gamma^3}{12} + \frac{5\gamma\rho}{16} + \frac{\gamma}{16} - \frac{11\xi}{40}\right)$$
$$\kappa_2 = 1 + \frac{1}{N}\left(\frac{\rho}{2} - \frac{13\gamma^2}{36}\right) + \frac{1}{N^2}\left(\gamma^2\left(-\frac{\gamma^2}{36} + \frac{53\rho}{48} + \frac{17}{48}\right) - \frac{251\gamma\xi}{120} - \frac{1}{2}\rho(\rho+1) + \frac{5\zeta}{3}\right),$$
$$\kappa_3 = \frac{1}{N^{3/2}}\left(-\frac{251\gamma^3}{216} + \frac{11\gamma\rho}{4} - \frac{9\xi}{5}\right),$$
$$\kappa_4 = \frac{1}{N^2}\left(-\frac{1079\gamma^4}{216} + \frac{33\gamma^2\rho}{2} - \frac{66\gamma\xi}{5} - \frac{35\rho^2}{6} + 8\zeta\right),$$
$$\kappa_5 = 0,$$
$$\kappa_6 = 0.$$

Cumulants to $\mathcal{O}(\mu^{-5/2})$ for the Poisson Model

Signed Wald statistic, R_W:

$\kappa_1 = \frac{1}{\mu^{3/2}}(-\gamma^3 + 2\gamma\rho - \xi)$,

$\kappa_2 = 1 + \frac{1}{\mu}(\rho - \gamma^2) + \frac{1}{\mu^2}(-10\gamma^4 + 30\gamma^2\rho - 22\gamma\xi - 8\rho^2 + 10\zeta)$,

$\kappa_3 = \frac{\gamma}{\sqrt{\mu}} + \frac{1}{\mu^{3/2}}(-12\gamma^3 + 21\gamma\rho - 9\xi)$,

$\kappa_4 = \frac{\rho}{\mu} + \frac{1}{\mu^2}(-114\gamma^4 + 288\gamma^2\rho - 180\gamma\xi - 60\rho^2 + 66\zeta)$,

$\kappa_5 = \frac{1}{\mu^{3/2}}(-30\gamma^3 + 50\gamma\rho - 19\xi)$,

$\kappa_6 = \frac{1}{\mu^2}(-114\gamma^4 + 288\gamma^2\rho - 180\gamma\xi - 60\rho^2 + 66\zeta)$.

Signed likelihood ratio statistic, R_{LR}:

$\kappa_1 = \frac{1}{\sqrt{\mu}}\left(-\frac{\gamma}{6}\right) + \frac{1}{\mu^{3/2}}\left(-\frac{\gamma^3}{12} + \frac{5\gamma\rho}{16} - \frac{11\xi}{40}\right)$

$\kappa_2 = 1 + \frac{1}{\mu}\left(\frac{\rho}{2} - \frac{13\gamma^2}{36}\right) + \frac{1}{\mu^2}\left(\gamma^2\left(\frac{53\rho}{48} - \frac{\gamma^2}{36}\right) - \frac{251\gamma\xi}{120} - \frac{\rho^2}{2} + \frac{5\zeta}{3}\right)$,

$\kappa_3 = \frac{1}{\mu^{3/2}}\left(-\frac{251\gamma^3}{216} + \frac{11\gamma\rho}{4} - \frac{9\xi}{5}\right)$,

$\kappa_4 = \frac{1}{\mu^2}\left(-\frac{1079\gamma^4}{216} + \frac{33\gamma^2\rho}{2} - \frac{66\gamma\xi}{5} - \frac{35\rho^2}{6} + 8\zeta\right)$, $\kappa_5 = 0$, $\kappa_6 = 0$.

Signed score statistic, R_S:

$\kappa_1 = 0$, $\kappa_2 = 1$, $\kappa_3 = \frac{\gamma}{\sqrt{\mu}}$, $\kappa_4 = \frac{\rho}{\mu}$, $\kappa_5 = \frac{\xi}{\mu^{3/2}}$, $\kappa_6 = \frac{\zeta}{\mu^2}$.

Cumulants to $\mathcal{O}(\mu^{-5/2})$ for the Poisson Model (cont'd)

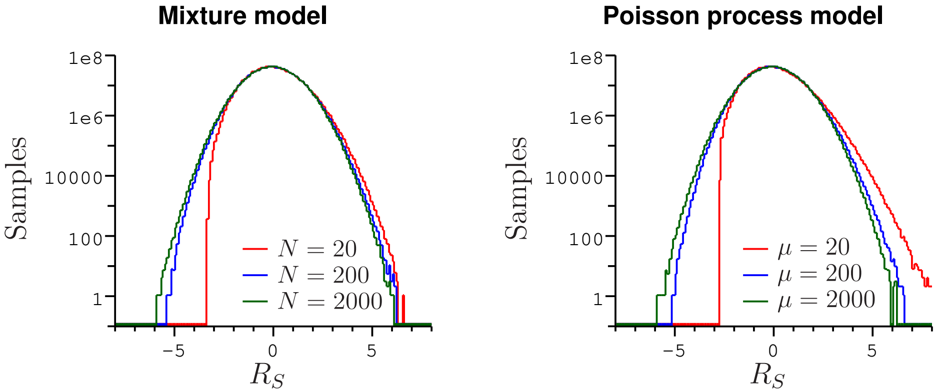
Signed Wald statistic using observed information, R_{W2}:

$$\kappa_1 = \frac{1}{\sqrt{\mu}}\left(-\frac{\gamma}{2}\right) + \frac{1}{\mu^{3/2}}\left(\frac{\gamma^3}{4} + \frac{5\gamma\rho}{16} - \frac{9\xi}{8}\right)$$
$$\kappa_2 = 1 + \frac{1}{\mu}\left(3\rho - \frac{5\gamma^2}{4}\right) + \frac{1}{\mu^2}\left(\frac{41\gamma^4}{4} - \frac{443\gamma^2\rho}{16} + \frac{55\gamma\xi}{8} + 10\rho^2 + 6\zeta\right)$$
$$\kappa_3 = \frac{1}{\sqrt{\mu}}(-2\gamma) + \frac{1}{\mu^{3/2}}\left(\frac{25\gamma^3}{8} + \frac{3\gamma\rho}{4} - 15\xi\right)$$
$$\kappa_4 = \frac{1}{\mu}(10\rho) + \frac{1}{\mu^2}\left(\frac{243\gamma^4}{8} - 132\gamma^2\rho + 45\gamma\xi + 72\rho^2 + 88\zeta\right)$$
$$\kappa_5 = \frac{1}{\mu^{3/2}}(15\gamma^3 - 40\gamma\rho - 54\xi)$$
$$\kappa_6 = \frac{1}{\mu^2}(-60\gamma^4 - 150\gamma^2\rho + 360\rho^2 + 330\gamma\xi + 376\zeta)$$

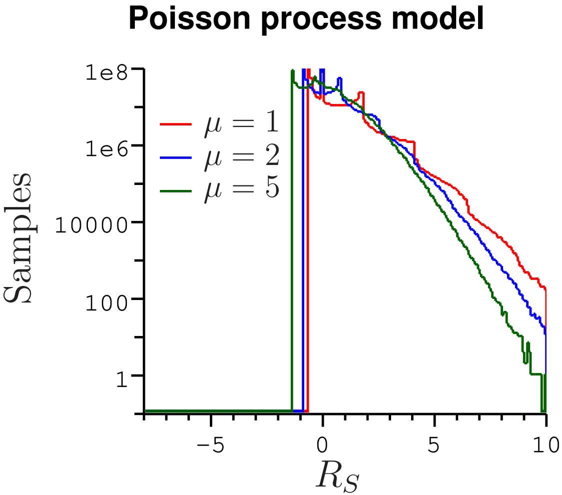
Signed Wald statistic using $\Delta\ell(\alpha) = -1/2$ uncertainty, R_{W3}:

$$\kappa_1 = \frac{1}{\sqrt{\mu}}\left(-\frac{\gamma}{2}\right) + \frac{1}{\mu^{3/2}}\left(-\frac{4\gamma^3}{9} + \frac{259\gamma\rho}{144} - \frac{15\xi}{8}\right)$$
$$\kappa_2 = 1 + \frac{1}{\mu}\left(\frac{7\rho}{2} - \frac{65\gamma^2}{36}\right) + \frac{1}{\mu^2}\left(\frac{170\gamma^4}{27} - \frac{583\gamma^2\rho}{48} - \frac{259\gamma\xi}{40} + \frac{131\rho^2}{36} + \frac{124\zeta}{9}\right)$$
$$\kappa_3 = \frac{1}{\sqrt{\mu}}(-2\gamma) + \frac{1}{\mu^{3/2}}\left(-\frac{5\gamma^3}{24} + \frac{107\gamma\rho}{12} - \frac{39\xi}{2}\right)$$
$$\kappa_4 = \frac{1}{\mu}(10\rho) + \frac{1}{\mu^2}\left(\frac{163\gamma^4}{8} - \frac{748\gamma^2\rho}{9} - 4\gamma\xi + \frac{130\rho^2}{3} + 124\zeta\right)$$
$$\kappa_5 = \frac{1}{\mu^{3/2}}(15\gamma^3 - 40\gamma\rho - 54\xi)$$
$$\kappa_6 = \frac{1}{\mu^2}(-60\gamma^4 - 150\gamma^2\rho + 360\rho^2 + 330\gamma\xi + 376\zeta)$$

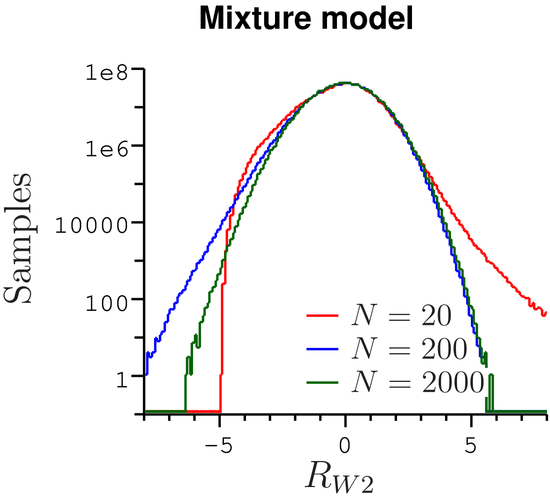
Simulated Distributions of R_S



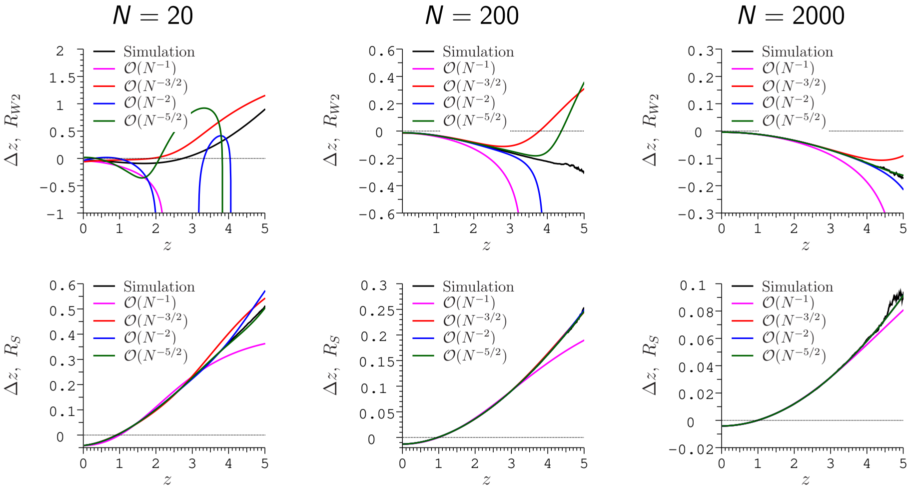
Simulated Distributions of R_s for Small μ



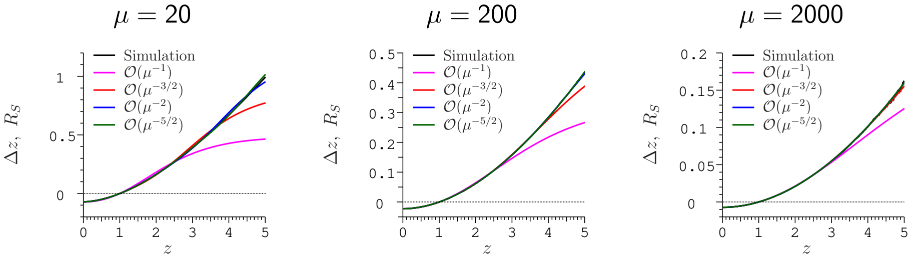
Simulated Distributions of R_{W2}



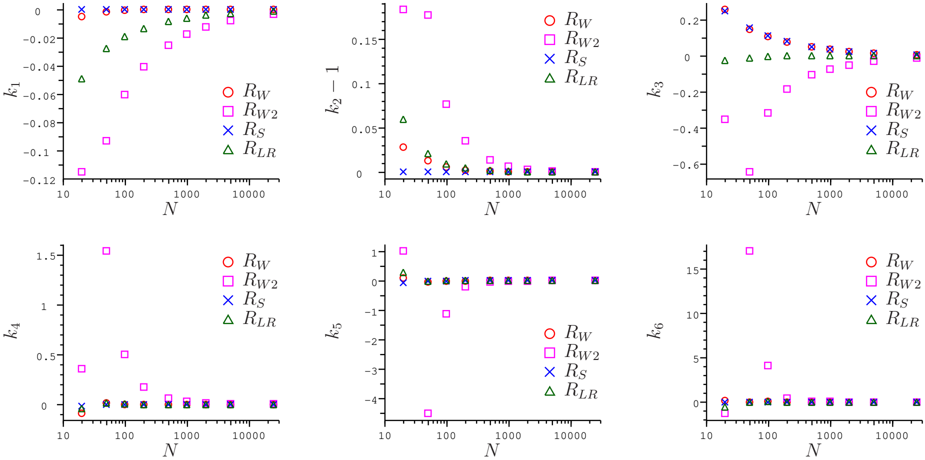
z Errors, Mixture Model



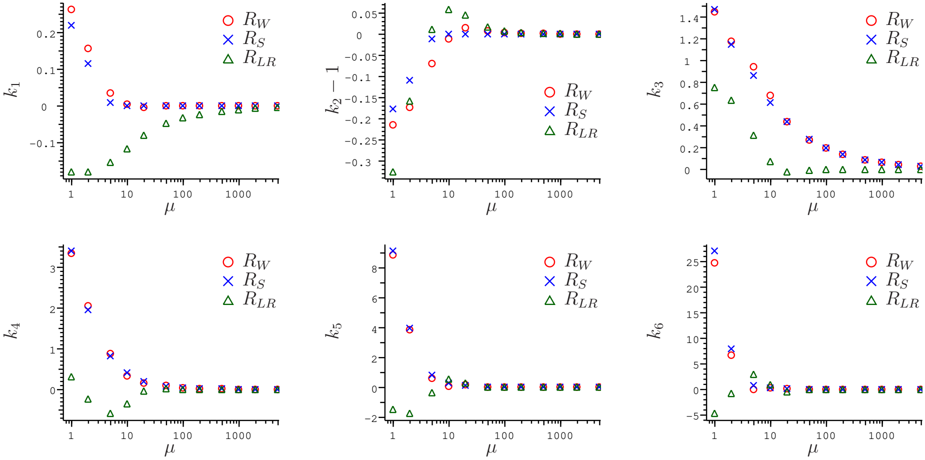
z Errors for R_S, Poisson Process Model



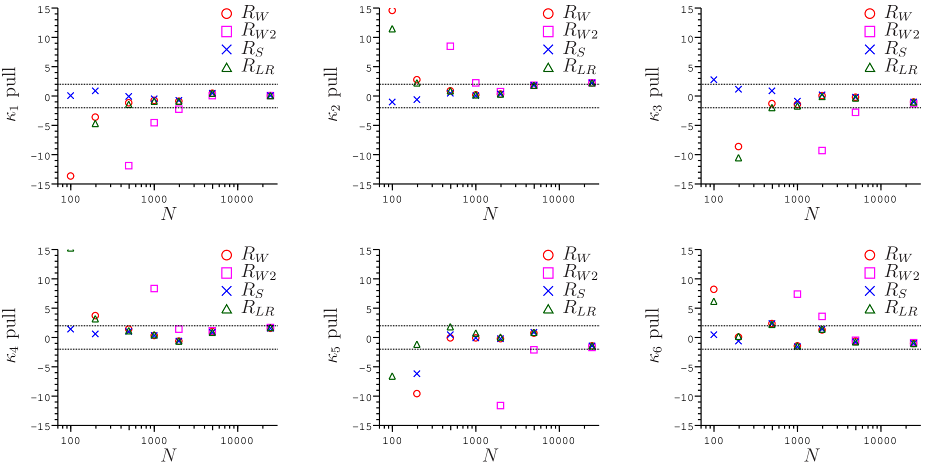
k-Statistics of Simulated Distributions, Mixture Model



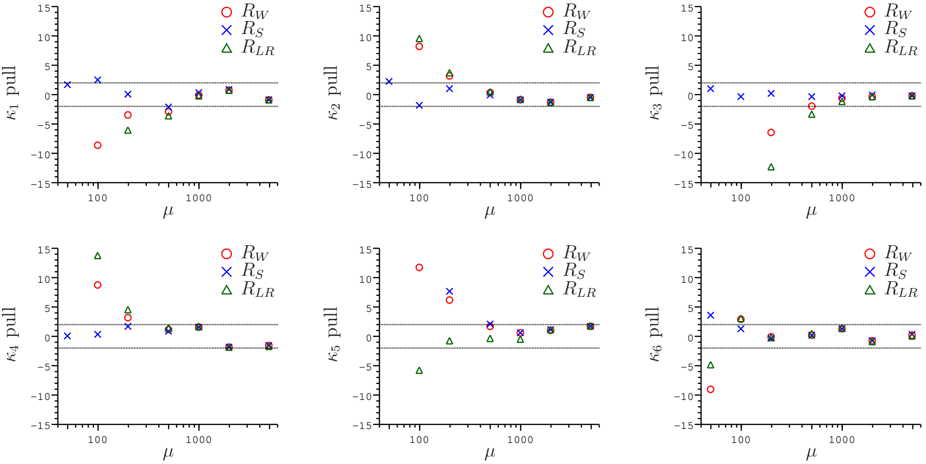
k-Statistics of Simulated Distributions, Poisson Model



Cumulant Matching to $\mathcal{O}(N^{-3/2})$, Mixture Model

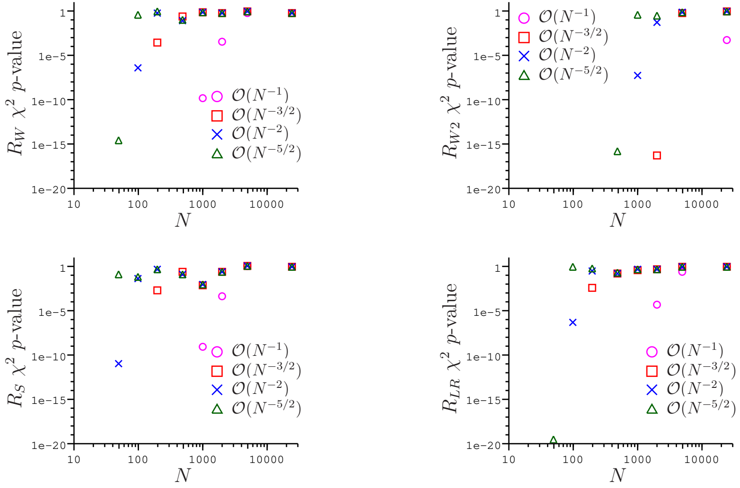


Cumulant Matching to $\mathcal{O}(\mu^{-3/2})$, Poisson Process Model



χ^2 Test for $z > 0$, Mixture Model

χ^2 is calculated for $z > 0$ bins of width 0.1 with 25 or more predicted counts, typically resulting in slightly over 50 bins used. This test is not affected by the lower limit on $\hat{\alpha}$ but it is still affected by the upper limit.



χ^2 is calculated for $z > 0$ bins of width 0.1 with 25 or more predicted counts.

The χ^2 p-values of the $\mathcal{N}(0, 1)$ approximation are below 10^{-20} for all statistics and for all N and μ values considered on this and previous slide.