Multiplicities in Au-Au and Cu-Cu collisions at $\sqrt{s_{NN}} = 62.4$ and 200 GeV

Dariusz Prorok

Institute of Theoretical Physics
University of Wrocław

Wrocław
January 7, 2013

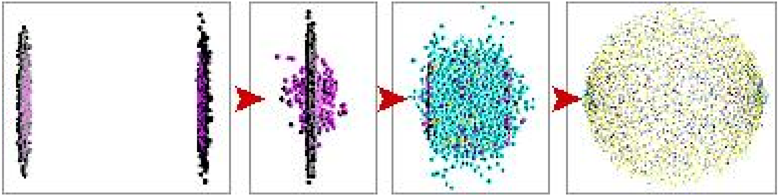
Sources

DP, arXiv:1208.0496

Data:

PHENIX Collaboration, Phys. Rev. C **78**, 044902 (2008)

Schematic view of a nucleus-nucleus collision



$z = 0$ plane

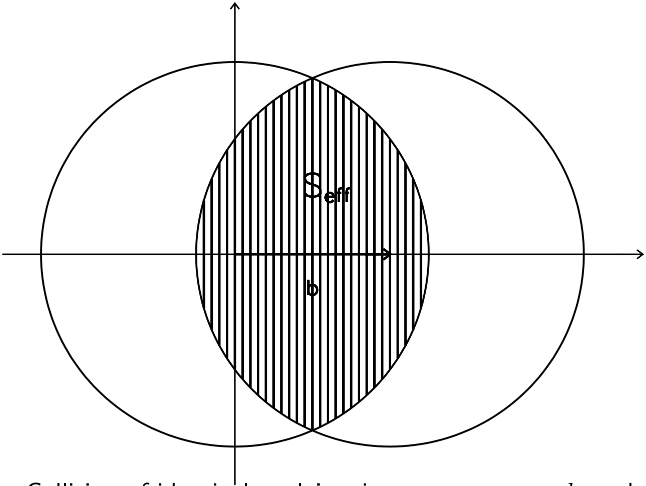
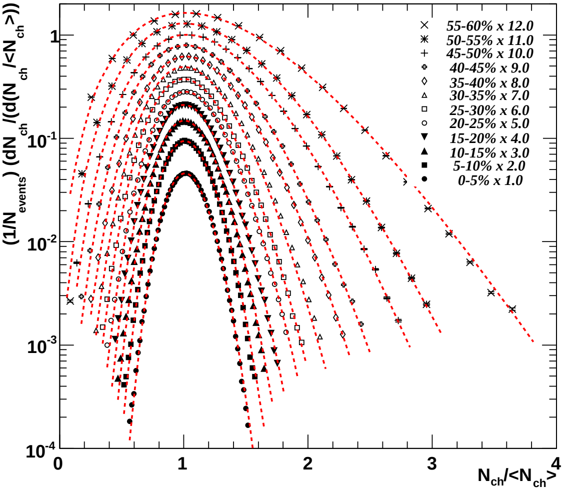


Figure: Collision of identical nuclei at impact parameter b at the moment of their maximal overlap.

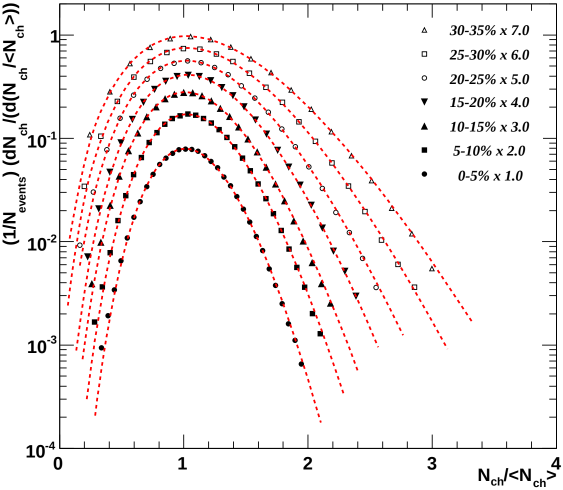
Multiplicity distribution: Au-Au at $\sqrt{s_{NN}} = 200$ GeV

$|\,\eta\,| < 0.26$



Multiplicity distribution: Cu-Cu at $\sqrt{s_{NN}} = 200$ GeV

$|\eta| < 0.26$



Negative binomial distribution

$$P(n;p,k) = \frac{k(k+1)(k+2)...(k+n-1)}{n!}(1-p)^n p^k$$

$0 \leq p \leq 1,$ k is a positive real number

$n = 0, 1, 2, ...$ - the number of charged particles in an event

$$\bar{n} = \frac{k(1-p)}{p}\,, \qquad V(n) = \frac{k(1-p)}{p^2}\,.$$

The maximum likelihood method

For N events in a sample there are N measurements of N_{ch}, say $\mathbf{X} = (X_1, X_2, ..., X_N)$.

$$L(\mathbf{X} \mid p, k) = \prod_{j=1}^{N} P(X_j; p, k)$$

The values $\hat{p}$ and $\hat{k}$ for which $L(\mathbf{X} \mid p, k)$ has its maximum are the maximum likelihood (ML) estimators of parameters p and k.

The log-likelihood function

$$\ln L(\mathbf{X} \mid p, k) = \sum_{j=1}^{N} \ln P(X_j; p, k)$$

Maximization of the log-likelihood function

$$\frac{\partial}{\partial p} \ln L(\mathbf{X} \mid p, k) = \sum_{j=1}^{N} \frac{\partial}{\partial p} \ln P(X_j; p, k) = 0$$

$$\frac{\partial}{\partial k} \ln L(\mathbf{X} \mid p, k) = \sum_{j=1}^{N} \frac{\partial}{\partial k} \ln P(X_j; p, k) = 0$$

For NBD the upper equation gives

$$\bar{n} = \langle N_{ch} \rangle \qquad \Longrightarrow \qquad \frac{1}{p} = \frac{\langle N_{ch} \rangle}{k} + 1$$

Likelihood ratio test - Wilks's theorem

X - a random variable with p.d.f $f(X, \theta)$, which depends on parameters $\theta = (\theta_1, \ \theta_2, ..., \theta_d) \in \Theta$, Θ is an open set in R^d.

$\mathbf{X} = (X_1, ..., X_N)$ - a sample of N independent observations of X

H_0 - a k-dimensional subset of Θ, $k < d$.

The maximum likelihood ratio:

$$\lambda = \frac{\max_{\theta \in H_0} L(\mathbf{X} \mid \theta)}{\max_{\theta \in \Theta} L(\mathbf{X} \mid \theta)}$$

If the hypothesis H_0 is true, *i.e.* it is true that $\theta \in H_0$, then the distribution of the statistic $-2 \ln \lambda$ converges to a χ^2 distribution with $d - k$ degrees of freedom as $N \longrightarrow \infty$.

χ^2 (chi-square) distribution

$0 \leq z \leq +\infty$,

$n = 1, 2, \ldots$ - the number of degrees of freedom

$$f(z;n) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1} \cdot e^{-z/2}$$

$$\Gamma(n) = (n-1)!\,, \quad \Gamma(x+1) = x\Gamma(x), \quad \Gamma(1/2) = \sqrt{\pi}$$

$$E[z] = n, \qquad V[z] = 2n$$

2-in-1 χ^2 function

Let define the function:

$$\chi^2(\mathbf{X} \mid \theta)_{\theta \in H_0} = -2 \ln \frac{L(\mathbf{X} \mid \theta)}{\max_{\theta' \in \Theta} L(\mathbf{X} \mid \theta')}$$

- The minimum of χ^2 with respect to $\theta \in H_0$ is at $\hat{\theta}$ - the ML estimators.
- The test statistic $\chi^2_{min} = \chi^2(\mathbf{X} \mid \hat{\theta})$ has a χ^2 distribution in the large sample limit.

p-value of the test statistic

The probability of obtaining the value of the test statistic equal to or greater then the value just obtained for the present data set (*i.e.* χ^2_{min}), when repeating the whole experiment many times:

$$p = P(\chi^2 \geq \chi^2_{min}; n_{dof}) = \int_{\chi^2_{min}}^{\infty} f(z; n_{dof}) dz \ ,$$

$f(z; n_{dof})$ - the χ^2 p.d.f.

$n_{dof} = d - k$ - the number of degrees of freedom

2-in-1 χ^2 function for binned data

Let divide the sample $\mathbf{X} = (X_1, X_2, ..., X_N)$ into m bins defined by the number of measured charged particles $\{0, 1, 2, 3, ..., m-1\}$ and with n_i entries in the ith bin, $N = \sum_{i=1}^{m} n_i$.

$$\chi^2 = -2 \ln \lambda = 2 \sum_{i=1}^{m} n_i \ln \frac{n_i}{\nu_i}$$

$$\nu_i = N \cdot P(i-1; p, k)$$

Details in: G. Cowan, *Statistical data analysis*, (Oxford University Press, Oxford, 1998)

2-in-1 χ^2 function for binned data, *cont.*

$$\chi^2(p,k) = 2\sum_{i=1}^{m} n_i \ln\frac{n_i}{\nu_i} = -2\,N\sum_{i=1}^{m} P_i^{ex}\ln\frac{P(i-1;p,k)}{P_i^{ex}}$$

$P_i^{ex} = n_i/N$ - the experimental probability (frequency)

- This χ^2 function depends explicitly on the number of events in the sample!
- But does not depend on actual experimental errors!

The χ^2 function of the least-squares method

The sum of squares of normalized residuals:

$$\chi^2_{LS}(p,k) = \sum_{i=1}^{m} \frac{(P_i^{ex} - P(i-1;p,k))^2}{err_i^2}$$

err_i - the uncertainty of the ith measurement

NOT MINIMIZED HERE !!!

but

$$\chi^2_{LS} = \chi^2_{LS}(\hat{p},\hat{k})$$

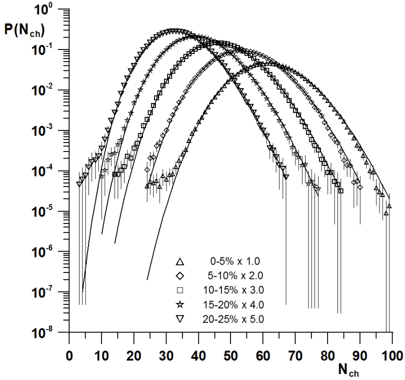
$\hat{p},\hat{k}$ - ML estimators of parameters p and k

Summary of motivations

- The fitted quantity is a probability distribution function (p.d.f.), so the most natural way is to use the maximum likelihood (ML) method, where the likelihood function is constructed directly from the tested p.d.f.. Because of Wilks's Theorem one can define a statistic, the distribution of which converges to a χ^2 distribution as the number of measurements goes to infinity. Thus for the large sample the goodness-of-fit can be expressed as a p-value computed with the corresponding χ^2 distribution.

- The most commonly used method, the least-squares method (LS) (called also χ^2 minimization), has the disadvantage of providing only the qualitative measure of the significance of the fit. Only if observables are represented by Gaussian random variables with known variances, the conclusion about the goodness-of-fit equivalent to that mentioned in the first point can be derived.

Multiplicity distributions: Au-Au at $\sqrt{s_{NN}} = 200$ GeV



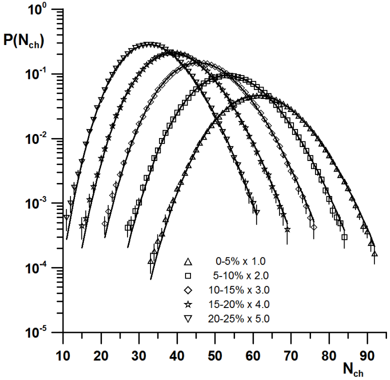


Figure: Uncorrected multiplicity distributions for bins with $n_i > 5$ (left) and $n_i > 60$ (right).

Multiplicity distributions: Au-Au at $\sqrt{s_{NN}} = 200$ GeV

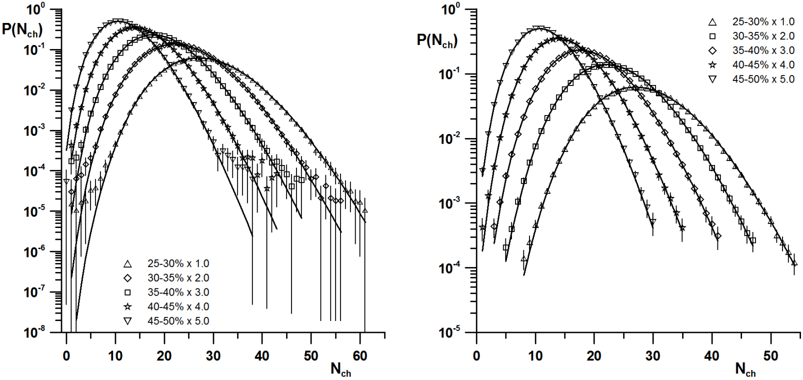


Figure: Uncorrected multiplicity distributions for bins with $n_i > 5$ (left) and $n_i > 60$ (right).

Results of fitting: Au-Au at $\sqrt{s_{NN}} = 200$ GeV, $n_i > 5$

Centr. %	N	$\hat{k}$	$\bar{n}$	χ^2/n_d	P %	χ^2_{LS}/n_d, err. quad.	stat.
0-5	653145	270.0± 2.5	61.85± 0.01	23.7	0	0.98	9.8
5-10	657944	163.4± 1.2	53.91± 0.01	9.1	0	0.69	6.9
10-15	658739	112.5± 0.7	46.50± 0.01	11.5	0	0.66	6.6
15-20	659607	85.1± 0.5	39.72± 0.01	8.9	0	0.52	5.2
20-25	658785	67.6± 0.4	33.56± 0.01	13.5	0	0.46	4.6
25-30	659632	56.7± 0.3	28.01± 0.01	10.9	0	0.37	3.7
30-35	659303	47.4± 0.3	23.02± 0.01	7.9	0	0.31	3.1
35-40	661174	40.5± 0.2	18.64± 0.01	8.5	0	0.37	3.7
40-45	661599	34.0± 0.2	14.84± 0.01	7.3	0	0.35	3.5
45-50	661765	27.3± 0.2	11.57± 0.005	10.5	0	0.92	9.2
50-55	662114	21.3± 0.1	8.82± 0.004	38.8	0	12.06	120.6

Results of fitting: Au-Au at $\sqrt{s_{NN}} = 200$ GeV, $n_i > 60$

Centr. %	N	$\hat{k}$	$\bar{n}$	χ^2/n_d	P %	χ^2_{LS}/n_d, err. quad.	stat.
0-5	652579	289.0± 2.9	61.86± 0.01	20.0	0	0.57	5.7
5-10	657571	168.1± 1.2	53.91± 0.01	20.6	0	0.61	6.1
10-15	658258	116.4± 0.7	46.50± 0.01	18.4	0	0.53	5.3
15-20	659302	86.9± 0.5	39.72± 0.01	12.6	0	0.43	4.3
20-25	658461	69.1± 0.4	33.56± 0.01	12.3	0	0.34	3.4
25-30	659337	57.9± 0.3	28.0± 0.01	10.4	0	0.28	2.8
30-35	659021	48.3± 0.3	23.02± 0.01	8.6	0	0.16	1.6
35-40	660937	41.3± 0.2	18.64± 0.01	7.6	0	0.19	1.9
40-45	661422	34.6± 0.2	14.84± 0.01	7.9	0	0.21	2.1
45-50	661577	27.9± 0.2	11.56± 0.005	10.0	0	0.23	2.3
50-55	661877	21.9± 0.1	8.81± 0.004	40.0	0	0.30	3.0

Multiplicity distributions: Au-Au at $\sqrt{s_{NN}} = 62.4$ GeV

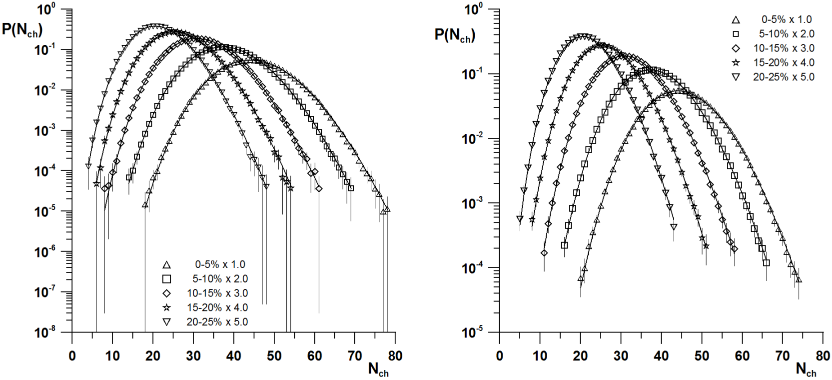


Figure: Uncorrected multiplicity distributions for bins with $n_i > 5$ (left) and $n_i > 60$ (right).

Multiplicity distributions: Au-Au at $\sqrt{s_{NN}} = 62.4$ GeV

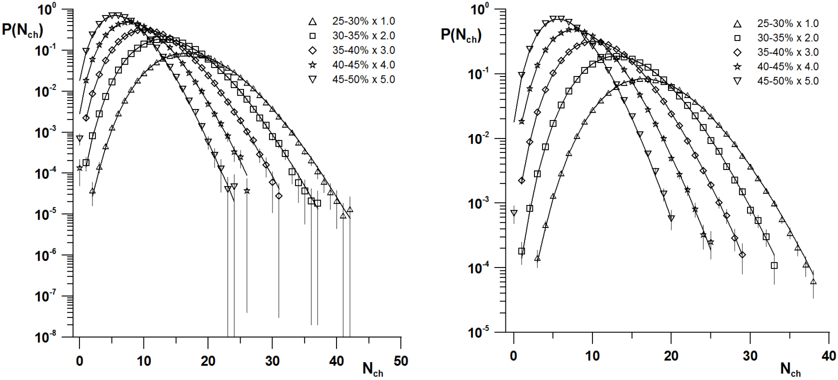


Figure: Uncorrected multiplicity distributions for bins with $n_i > 5$ (left) and $n_i > 60$ (right).

Results of fitting: Au-Au at $\sqrt{s_{NN}} = 62.4$ GeV, $n_i > 5$

Centr. %	N	$\hat{k}$	$\bar{n}$	$\frac{\chi^2}{n_d}$	P %	χ^2_{LS}/n_d, err. quad.	χ^2_{LS}/n_d, err. stat.
0-5	607155	225.2$\pm$2.5	44.67$\pm$0.01	2.4	10^{-8}	0.18	1.8
5-10	752392	142.3$\pm$1.1	37.96$\pm$0.01	2.4	10^{-8}	0.11	1.1
10-15	752837	115.2$\pm$0.9	31.53$\pm$0.01	2.1	10^{-5}	0.13	1.3
15-20	752553	88.0$\pm$0.6	26.07$\pm$0.01	1.9	10^{-4}	0.13	1.3
20-25	752296	68.5$\pm$0.5	21.35$\pm$0.01	2.6	10^{-8}	0.21	2.1
25-30	752183	53.2$\pm$0.4	17.30$\pm$0.01	2.7	10^{-8}	0.23	2.3
30-35	751375	40.1$\pm$0.3	13.84$\pm$0.005	3.0	10^{-8}	0.25	2.5
35-40	751661	31.7$\pm$0.2	10.89$\pm$0.004	6.7	0	0.16	1.6
40-45	750884	25.1$\pm$0.2	8.42$\pm$0.004	37.5	0	40.36	403.6
45-50	751421	21.8$\pm$0.2	6.41$\pm$0.003	209	0	285.9	2859

Results of fitting: Au-Au at $\sqrt{s_{NN}} = 62.4$ GeV, $n_i > 60$

Centr.					P	χ^2_{LS}/n_d, err.	
%	N	$\hat{k}$	$\bar{n}$	χ^2/n_d	%	quad.	stat.
0-5	607075	227.9 ± 2.5	44.67 ± 0.01	5.6	0	0.19	1.9
5-10	752263	143.9 ± 1.1	37.96 ± 0.01	7.8	0	0.12	1.2
10-15	752739	116.2 ± 0.9	31.53 ± 0.01	5.7	0	0.13	1.3
15-20	752492	88.5 ± 0.6	26.07 ± 0.01	6.0	0	0.11	1.1
20-25	752182	69.2 ± 0.5	21.35 ± 0.01	10.2	0	0.22	2.2
25-30	752095	53.6 ± 0.4	17.30 ± 0.01	8.2	0	0.23	2.3
30-35	751324	40.3 ± 0.3	13.84 ± 0.005	7.4	0	0.26	2.6
35-40	751639	31.8 ± 0.2	10.89 ± 0.004	9.4	0	0.15	1.5
40-45	750852	25.2 ± 0.2	8.42 ± 0.004	51	0	0.22	2.2
45-50	751348	22.0 ± 0.2	6.41 ± 0.003	260	0	343	3431

Multiplicity distributions: Cu-Cu at $\sqrt{s_{NN}} = 200$ GeV

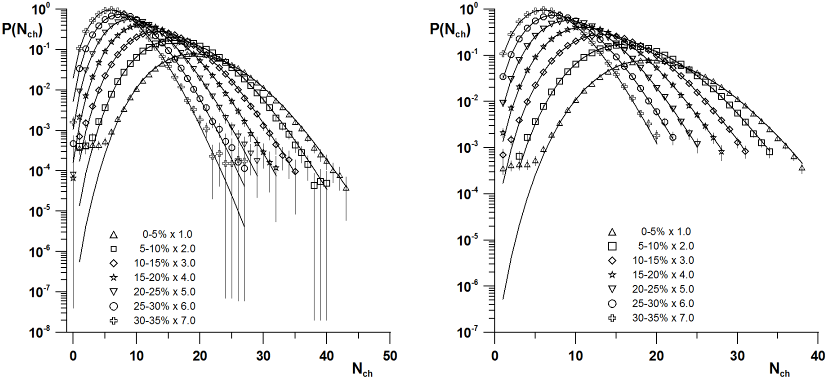


Figure: Uncorrected multiplicity distributions for bins with $n_i > 5$ (left) and $n_i > 60$ (right).

Results of fitting: Cu-Cu at $\sqrt{s_{NN}} = 200$ GeV, $n_i > 60$

Centr. %	N	$\hat{k}$	$\bar{n}$	χ^2/n_d	P %	χ^2_{LS}/n_d, err. quad.	stat.
0-5	368271	61.5 ± 0.6	19.79 ± 0.01	122.2	0	2.3	23.0
5-10	368869	52.0 ± 0.5	16.74 ± 0.01	20.5	0	0.39	3.9
10-15	369825	42.3 ± 0.4	14.05 ± 0.01	16.2	0	0.43	4.3
15-20	369964	35.1 ± 0.3	11.77 ± 0.01	11.4	0	0.24	2.4
20-25	371752	29.8 ± 0.3	9.80 ± 0.01	6.6	0	0.20	2.0
25-30	368708	25.6 ± 0.3	8.14 ± 0.01	42.7	0	0.21	2.1
30-35	367869	22.6 ± 0.2	6.72 ± 0.005	126.4	0	0.62	6.2

Multiplicity distributions: Cu-Cu at $\sqrt{s_{NN}} = 62.4$ GeV

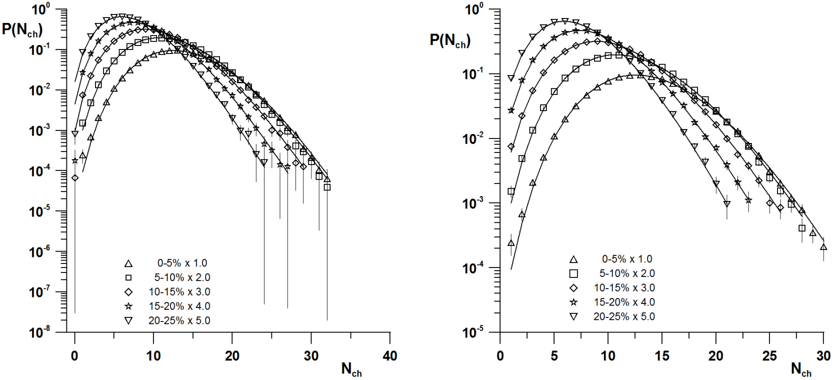


Figure: Uncorrected multiplicity distributions for bins with $n_i > 5$ (left) and $n_i > 60$ (right).

Results of fitting: Cu-Cu at $\sqrt{s_{NN}} = 62.4$ GeV, $n_i > 60$

Centr.					P	χ^2_{LS}/n_d, err.	
%	N	$\hat{k}$	$\bar{n}$	χ^2/n_d	%	quad.	stat.
0-5	298131	42.0 ± 0.5	13.35 ± 0.01	14.7	0	0.67	6.7
5-10	307061	26.8 ± 0.2	11.66 ± 0.01	19.7	0	0.86	8.6
10-15	309798	20.7 ± 0.2	9.90 ± 0.01	19.4	0	0.38	3.8
15-20	312434	18.0 ± 0.1	8.27 ± 0.01	46.5	0	0.40	4.0
20-25	312758	16.3 ± 0.1	6.89 ± 0.01	118.1	0	0.63	6.3

Conclusions

1. Results of the likelihood ratio tests suggest that the hypothesis about the NBD of charged-particle multiplicities measured by the PHENIX Collaboration should be rejected.
2. The significant systematic errors of the data are the reasons for acceptable values of the LS test statistic for almost all centrality classes of PHENIX measurements.
3. The size of the sample is very important for the validation of a hypothesis about the p.d.f. of an observable. If the hypothesis is true, the distribution is exact for the whole population. Thus for the very large samples the measured distribution should be very close to that postulated. How "close" should be controlled by the size of the sample, i.e., discrepancies should decrease as N_{event} increases.

The least-squares method - the so-called χ^2 minimization

$$\chi^2(\alpha_1, ..., \alpha_l) = \sum_{k=1}^{n} \frac{\left[R_k^{exp} - R_k^{th}(\alpha_1, ..., \alpha_l)\right]^2}{\sigma_k^2}$$

$$n_{dof} = n - l$$

$\hat{\alpha}_1, ..., \hat{\alpha}_l$ - the LS estimators of $\alpha_1, ..., \alpha_l$ = the values of these parameters at the minimum of χ^2

The key test number of LS method:

$$\boxed{\frac{\chi^2(\hat{\alpha}_1, ..., \hat{\alpha}_l)}{n_{dof}} \sim 1}$$