Interpolazione della distribuzione di massa invariante (segnale di una particella)

Esercitazione del Corso di *Laboratorio Analisi Dati*

Secondo anno Magistrale / Primo Semestre

docente: **Alexis Pompili**

Preliminarmente visualizziamo la distribuzione che deve essere interpolata:

root [0] .L *d0kpi-massfit.C*

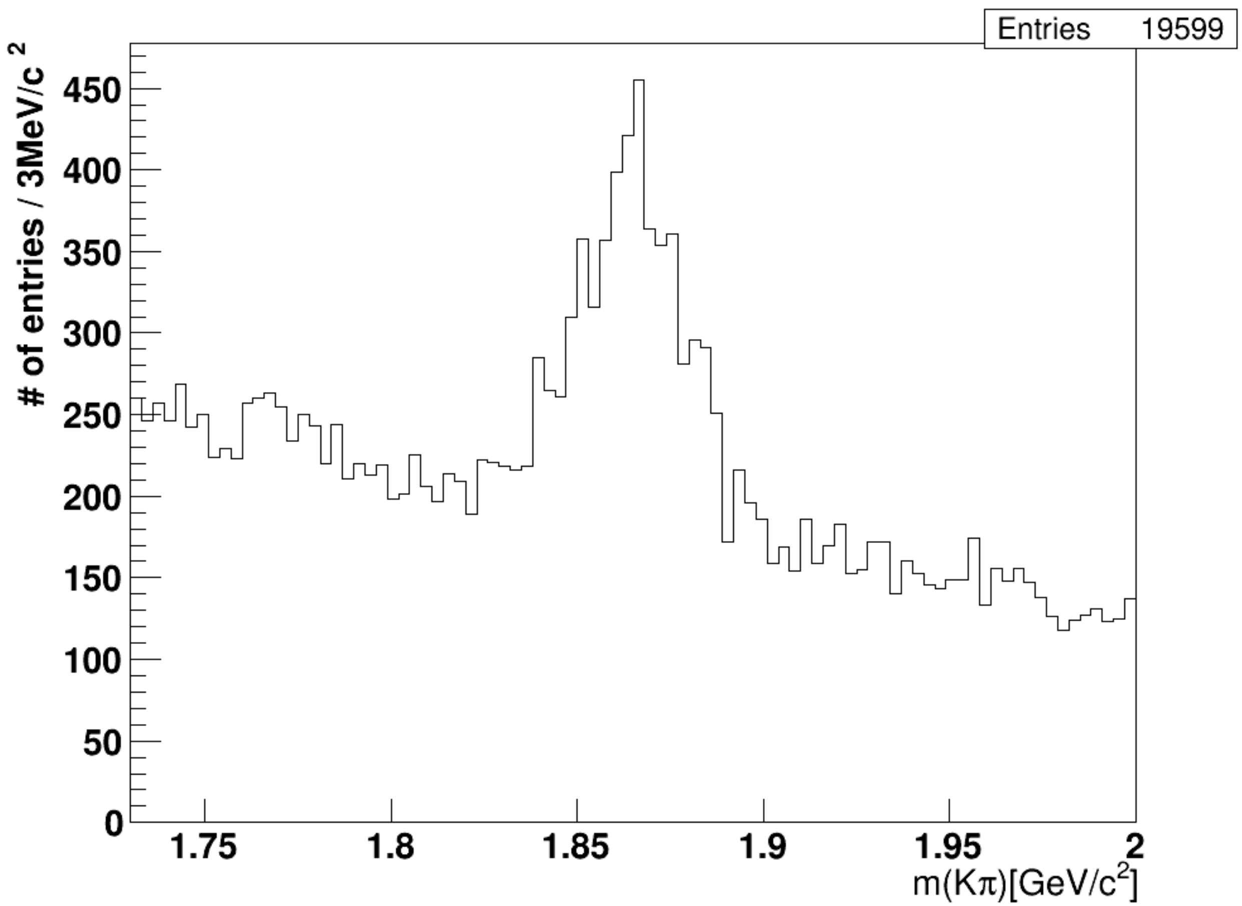
root [1] *main("17dec","png","Step0")*

Macro file (C^{++} program)

Passo-0 ed esecuzione relativa parte di codice

data

formato plot di output



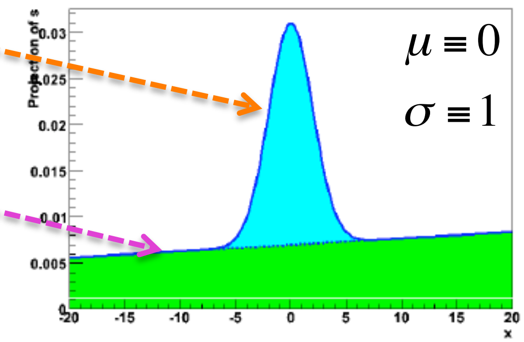
Preliminarmente discutiamo il modello di interpolazione che pensiamo di usare al fine di implementare la relativa PDF :

1) **denoto** $x \equiv m_{K\pi}$

2) $f(x) = f_{sig}(x) + f_{bkg}(x)$ **non e' una PDF: non e' normalizzata poiche'**

$$f_{sig}(x) = \frac{A}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$ **e' una funzione gaussiana**

$$f_{bkg}(x) = a + bx$$ **e' una retta (polinomiale di ord.1)**

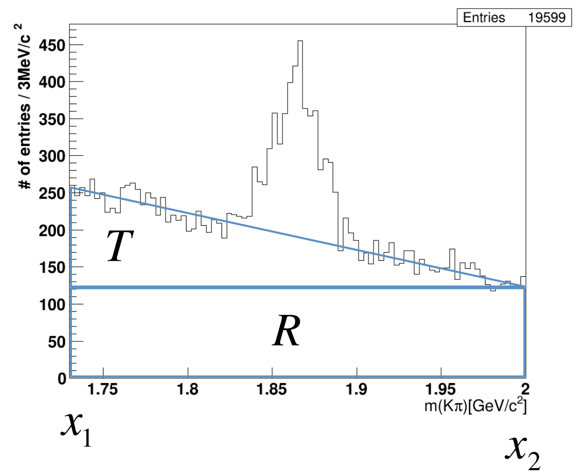


3) **come si normalizzano separatamente le due funzioni?**

$$\tilde{f}_{sig}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$ **(banalmente si elimina il coeff. moltiplicativo A)**

$$\tilde{f}_{bkg}(x) = \frac{a+bx}{\int_{x_1}^{x_2}(a+bx)\,dx}$$ **(si divide per l'area sotto la retta)**

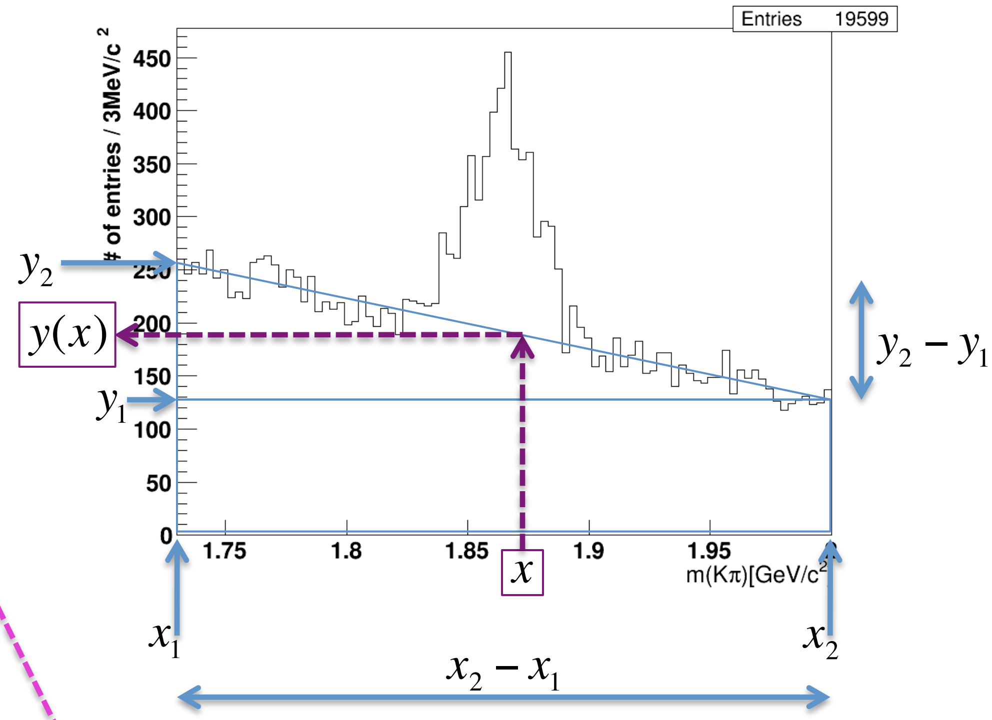
$$a(x_2 - x_1) + \frac{1}{2}b(x_2^2 - x_1^2) = T + R$$



dimostrazione alla ... slide seguente !

Dimostrazione:

equazione della retta: $\underbrace{y_1 - y_2}_{<0} = \underbrace{b}_{<0}\underbrace{(x_2 - x_1)}_{>0}$



e in generale: $y_1 - y(x) = b(x_2 - x)$

che, se vogliamo nella forma $y(x) = a + bx$ **,**

richiede : $a = y_1 - bx_2$ **(cioe'** $y_1 = a + bx_2$ **)**

Calcoliamo ora le aree geometricamente:

- **Triangolo** $T = \frac{1}{2}(y_2 - y_1)(x_2 - x_1)$
- **Rettangolo** $R = y_1(x_2 - x_1)$

L'area totale …
sotto la retta e': $R + T = y_1(x_2 - x_1) + \frac{1}{2}(y_2 - y_1)(x_2 - x_1) = (a + bx_2)(x_2 - x_1) + \frac{1}{2}b(x_1 - x_2)(x_2 - x_1)$

da cui, con un po' di algebra, si ottiene: $\boxed{T + R = a(x_2 - x_1) + \frac{1}{2}b(x_2^2 - x_1^2)}$ **(c.v.d.)**

La PDF complessiva deve essere a sua volta normalizzata, per cui non puo' essere banalmente $\tilde{f}(x) = \tilde{f}_{sig}(x) + \tilde{f}_{bkg}(x)$ **bensi':** $\tilde{f}(x) = f_S \tilde{f}_{sig}(x) + (1 - f_S)\tilde{f}_{bkg}(x)$

… essendo $f_S = \dfrac{N_{sig}}{N_{tot}} = \dfrac{\#\,\text{candidati di segnale}}{\#\,\text{candidati totale}}$ **la cosidetta** ***frazione di segnale*****.**

Ma questa PDF non puo' interpolare l'istogramma (avendo appunto area unitaria) ed e' quindi necessario introdurre un ulteriore parametro moltiplicativo che consenta di scalare la funzione PDF (normalizzata) all'istogramma:

$$f_{FIT}(x) = c_s \cdot \tilde{f}(x)$$

dove c_s **e' il cosidetto** ***fattore di scala*****.**

Si noti che i parametri liberi sono 6 e non 7 come potrebbe sembrare poiche' N_{tot} **viene fissato al numero totale di entrate nell'istogramma!**

Pertanto il modello di interpolazione (a 6 parametri liberi) e' dato dalla:

$$f_{FIT}(x\ ;\ \underbrace{\mu,\sigma}_{\text{del Segnale}},\ \underbrace{a,b}_{\text{del Fondo}},\ \underbrace{N_S, c_s}_{\text{comuni a Segnale e Fondo}})$$

Per eseguire l'interpolazione (STEP-1):

root [0] *.L d0kpi-massfit.C*
root [1] *main("17dec","png","Step1")*

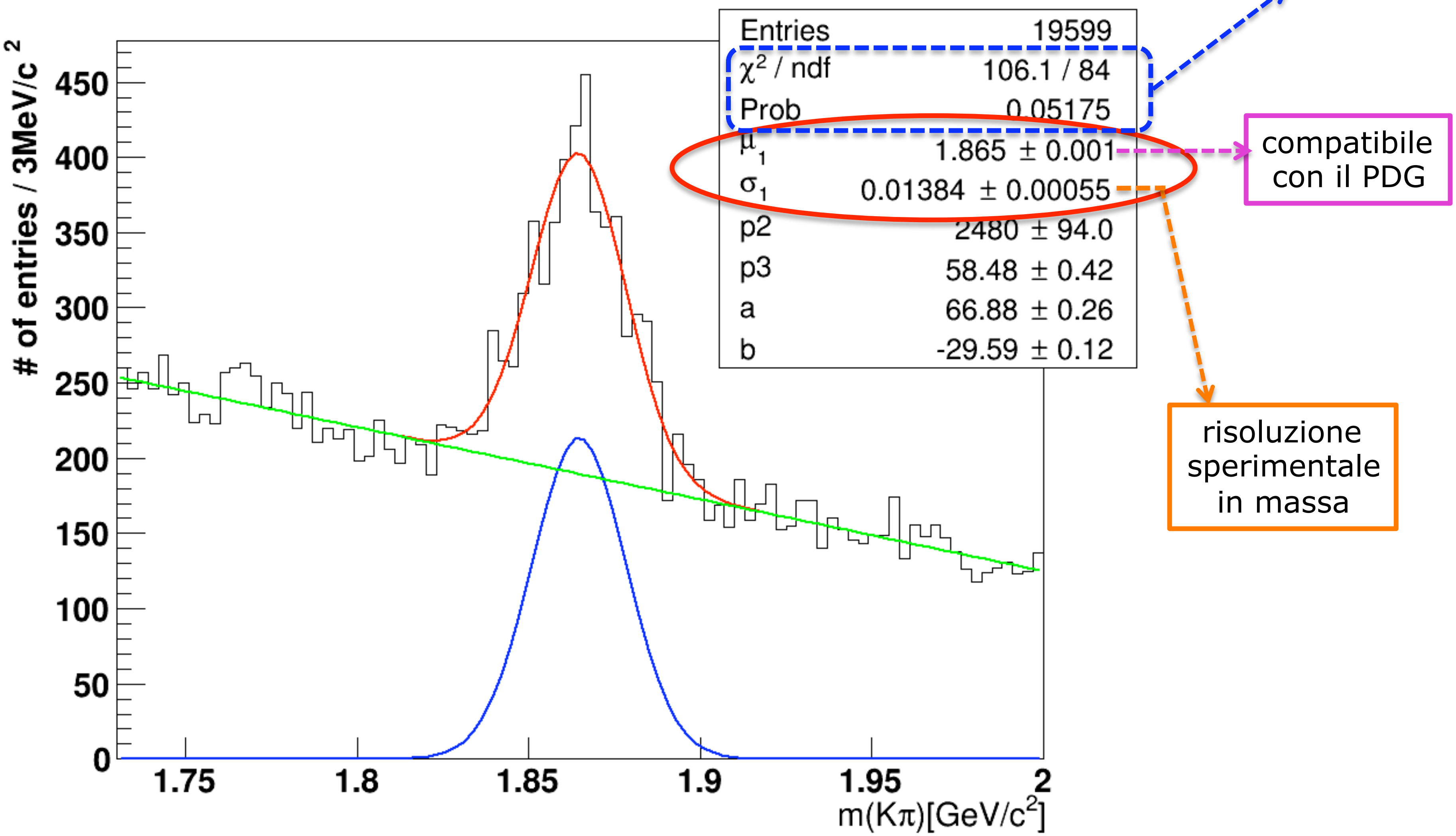
Ottenendo …

```
FCN=106.142 FROM MIGRAD    STATUS=CONVERGED     131 CALLS        132 TOTAL
                     EDM=1.18784e-08    STRATEGY= 1  ERROR MATRIX UNCERTAINTY   4.9 per cent
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
   1  #mu_{1}      1.86479e+00   5.52379e-04  -1.70695e-07  -1.96244e-01
   2  #sigma_{1}   1.38370e-02   5.52320e-04   3.21210e-07  -1.51908e-01
   3  p2           2.48036e+03   9.40077e+01  -1.03036e-01   4.39934e-07
   4  p3           5.84786e+01   4.16854e-01   1.96901e-04   2.29055e-04
   5  p4           1.95990e+04     fixed
   6  a            6.68809e+01   2.56167e-01  -1.92659e-05  -3.40888e-04
   7  b           -2.95914e+01   1.20525e-01  -9.20883e-06  -7.68731e-04
```

$\boxed{\hat{m} \cong (1864.79 \pm 0.55) MeV}$

$\boxed{\hat{\sigma} \cong (13.84 \pm 0.55) MeV}$

Risultato dell'interpolazione con un modello
che prevede una singola gaussiana per il segnale
ed una polinomiale di ord.1 per il fondo (fondo lineare) :



Argomentiamo le considerazioni precedenti :

1) Massa : J. Beringer *et al.* (Particle Data Group), PR **D86**, 010001 (2012)

D^0 $\qquad I(J^P) = \frac{1}{2}(0^-)$

D^0 MASS

The fit includes $D^{\pm}$, D^0, $D_s^{\pm}$, $D^{*\pm}$, D^{*0}, $D_s^{*\pm}$, $D_1(2420)^0$, $D_2^*(2460)^0$, and $D_{s1}(2536)^{\pm}$ mass and mass difference measurements.

VALUE (MeV)	EVTS	DOCUMENT ID		TECN	COMMENT
1864.86 ± 0.13 OUR FIT					
1864.91 ± 0.17 OUR AVERAGE					
1865.30 ± 0.33 ±0.23	98 ± 13	ANASHIN	10A	KEDR	e^+e^- at $\psi(3770)$
1864.847 ± 0.150 ± 0.095	319 ± 18	CAWLFIELD	07	CLEO	$D^0 \to K_S^0 \phi$
1864.6 ± 0.3 ±1.0	641	BARLAG	90C	ACCM	π^- Cu 230 GeV

2) Larghezza : BaBar Collaboration, Phys. Rev. D 88, 052003 (2013)

We measure the mass difference, Δm_0, between the $D^*(2010)^+$ and the D^0 and the natural line width, Γ, of the transition $D^*(2010)^+ \to D^0\pi^+$. The data were recorded with the *BABAR* detector at center-of-mass energies at and near the $\Upsilon(4S)$ resonance, and correspond to an integrated luminosity of approximately 477 fb^{-1}. The D^0 is reconstructed in the decay modes $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^-\pi^+$. For the decay mode $D^0 \to K^-\pi^+$ we obtain $\Gamma = (83.4 \pm 1.7 \pm 1.5)$ keV and $\Delta m_0 = (145\,425.6 \pm 0.6 \pm 1.8)$ keV, where the quoted errors are statistical and systematic, respectively. For the $D^0 \to K^-\pi^+\pi^-\pi^+$ mode we obtain $\Gamma = (83.2 \pm 1.5 \pm 2.6)$ keV and $\Delta m_0 = (145\,426.6 \pm 0.5 \pm 2.0)$ keV. The combined measurements yield $\Gamma = (83.3 \pm 1.2 \pm 1.4)$ keV and $\Delta m_0 = (145\,425.9 \pm 0.4 \pm 1.7)$ keV; the width is a factor of approximately 12 times more precise than the previous value, while the mass difference is a factor of approximately 6 times more precise.

Interpolazione & calcolo di *yield*, significativita' statitica e purezza del segnale:

root [0] *.L d0kpi-massfit.C*
root [1] *main("17dec","png","Step2")*

```
 FCN=106.142 FROM MIGRAD    STATUS=CONVERGED     131 CALLS        132 TOTAL
                     EDM=1.18784e-08    STRATEGY= 1  ERROR MATRIX UNCERTAINTY   4.9 per cent
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
   1  #mu_{1}      1.86479e+00   5.52379e-04  -1.70695e-07  -1.96244e-01  --------> massa
   2  #sigma_{1}   1.38370e-02   5.52320e-04   3.21210e-07  -1.51908e-01  --------> risoluzione
   3  p2           2.48036e+03   9.40077e+01  -1.03036e-01   4.39934e-07  --------> candidati di segnale
   4  p3           5.84786e+01   4.16854e-01   1.96901e-04   2.29055e-04  --------> fattore di scala
   5  p4           1.95990e+04     fixed
   6  a            6.68809e+01   2.56167e-01  -1.92659e-05  -3.40888e-04  --------> intercetta
   7  b           -2.95914e+01   1.20525e-01  -9.20883e-06  -7.68731e-04  --------> coeff.ang.
chi2 = 106.142 , ndf = 84 so that: norm-chi2 = 1.26359
integral gauss (SIG-fraction) = 0.126555
integral background (BKG-fraction) = 0.873445
num. total candidates (as entries in the histogram) = 19599
num. signal candidates as estimated by the fit = 2480.36 +/- 94.0077
====== USING WHOLE FIT MODEL (S+B) ========
num. S+B  in (1.83712 , 1.89247) = 5878.56
num. S in (1.83712 , 1.89247) = 2367.5
num. B in (1.83712 , 1.89247) = 3511.06
====== USING ONLY BKG-FIT MODEL (BKG from sidebands) ========
num. S+B  in (1.835 , 1.895) = 6231
num. S in (1.835 , 1.895) = 2426.86
num. B in (1.835 , 1.895) = 3804.14
======= STATISTICAL SIGNIFICANCE & SIGNAL-TO-NOISE RATIO =======
 SS (CURVE) = 30.8784     S/B = 0.674299
 SS (HISTO) = 30.7444     S/B(istogramma) = 0.637951
```

…mentre il plot nella canvas:

In generale:

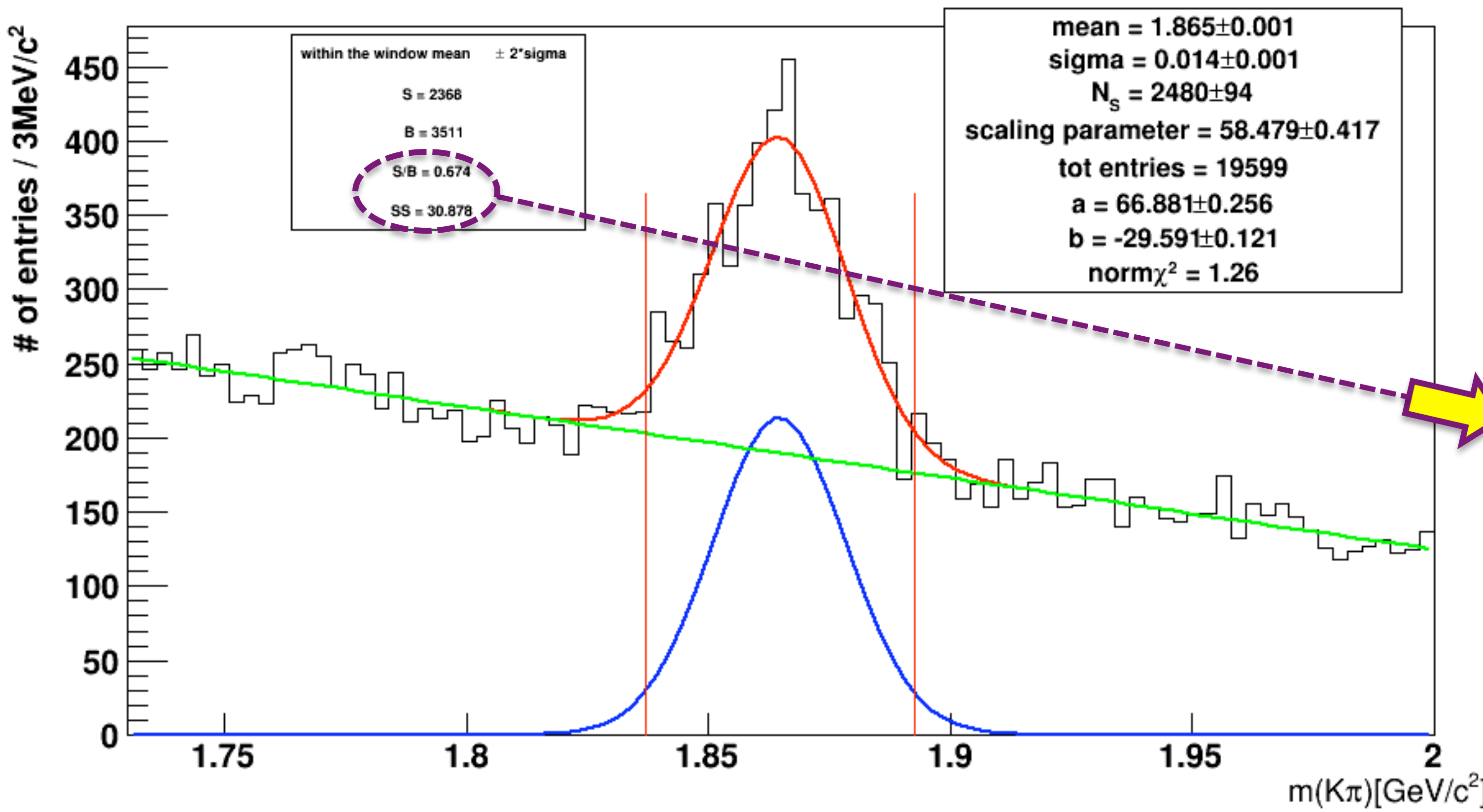
Significativita' statistica $SS = \frac{S^2}{S+B}$ **Purezza** $P = \frac{S}{B}$

Utilizzando l'intera informazione fornita dall'interpolazione :

$$SS_{fit} = \frac{S_{fit}}{(S+B)_{fit}} \qquad P_{fit} = \frac{S_{fit}}{B_{fit}}$$

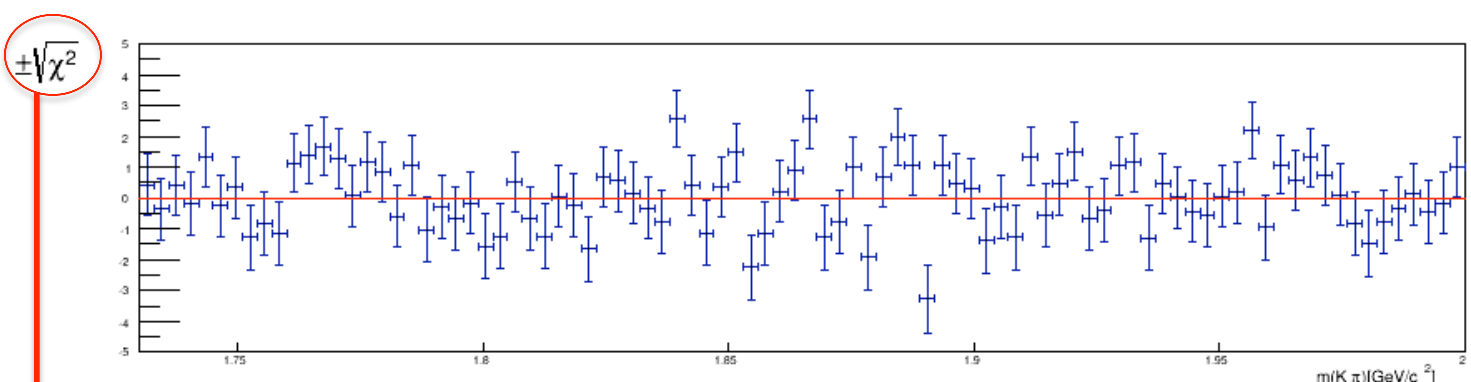
Utilizzando l'interpolazione parzialmente (per il solo fondo):

$$SS_{fit-histo} = \frac{S_{fit-histo}}{(S+B)_{histo}} = \frac{(S+B)_{histo} - B_{fit}}{(S+B)_{histo}} \qquad P_{fit-histo} = \frac{S_{fit-histo}}{B_{fit}} = \frac{(S+B)_{histo} - B_{fit}}{B_{fit}}$$



definizione alla … slide precedente !

Calcolo di significativita' statistica e di purezza.



Ulteriore analisi della bonta' del fit

definizione alla … slide seguente !

Definizione dello pseudo chi-quadrato o scarto normalizzato :

Lo scarto normalizzato e' simile alla radice del chi-quadrato corredato di segno (motivo per il quale tecnicamente e' uno "pseudo chi-quadrato"). Lo denoto con $\pm\sqrt{\chi^2}$

L'istogramma degli scarti normalizzati desidero che abbia lo stesso # di bin dell'istogramma della distribuzione di massa invariante.
E' necessario anche rappresentare lo scarto corredato dalla propria barra di errore!

Valore sperimentale

Valore atteso (teorico)

$$\pm\sqrt{\chi^2}(i) = \frac{x_S^i - x_T^i}{\sigma_i} \equiv \frac{N_i - F_i}{\sqrt{N_i}}$$

… essendo :

- N_i **= # candidati nel bin i-esimo dell'istogramma**
- F_i **= # candidati nel bin i-esimo atteso (assumendo corretto il modello di fit)**

Incertezza associata al valore sperimentale

L'incertezza (errore) sullo scarto normalizzato (per ogni bin) si calcola applicando l'usuale legge di propazazione degli errori casuali (tralascio l'indice per alleggerire la notazione):

$$\sigma^2_{\pm\sqrt{\chi^2}} = \left(\frac{d}{dN}\left(\frac{N-F}{\sqrt{N}}\right)\right)^2 \cdot \left(\sqrt{N}\right)^2 = \left(\frac{\sqrt{N} - \frac{1}{2\sqrt{N}}(N-F)}{N}\right)^2 \cdot N = \left(\frac{N - \frac{1}{2}(N-F)}{N\sqrt{N}}\right)^2 \cdot N = \left(\frac{1}{2}\left(\frac{N+F}{N}\right)\right)^2$$

In conclusione: $\sigma_{\pm\sqrt{\chi^2}} = \frac{1}{2}\frac{N+F}{N}$ **e, ad alta statistica (N grande) si ha:** $N \approx F \Rightarrow \sigma_{\pm\sqrt{\chi^2}} \approx 1$