Introducción a la probabilidad

The *cumulative distribution function* $F(a)$ is the probability that $x \le a$:

$$F(a) = \int_{-\infty}^{a} f(x)\, dx \ . \tag{32.6}$$

Here and below, if x is discrete-valued, the integral is replaced by a sum. The endpoint a is expressly included in the integral or sum. Then $0 \le F(x) \le 1$, $F(x)$ is nondecreasing, and $P(a < x \le b) = F(b) - F(a)$. If x is discrete, $F(x)$ is flat except at allowed values of x, where it has discontinuous jumps equal to $f(x)$.

Any function of random variables is itself a random variable, with (in general) a different p.d.f. The *expectation value* of any function $u(x)$ is

$$E[u(x)] = \int_{-\infty}^{\infty} u(x)\, f(x)\, dx \ , \tag{32.7}$$

assuming the integral is finite. For $u(x)$ and $v(x)$, any two functions of x, $E[u+v] = E[u] + E[v]$. For c and k constants, $E[cu+k] = cE[u] + k$.

The n^{th} moment of a random variable is

$$\alpha_n \equiv E[x^n] = \int_{-\infty}^{\infty} x^n f(x)\, dx \ , \tag{32.8a}$$

Variables aleatorias
Probabilidad

- Uno de los métodos consiste en utilizar la ley de los números grandes. En este caso, asumimos que podemos realizar cualquier número de tiros de la moneda, con cada tiro de la moneda siendo independiente — es decir, el resultado de cada tiro de la moneda no es afectado por los tiros anteriores de la moneda. Si realizamos *N* pruebas (tiros de la moneda), y dejamos N_H ser el número de veces que la moneda aterriza en “cabezas”, entonces nosotros podemos, para cualesquiera *N*, considerar el cociente N_H/N.
- Mientras que N se vuelve cada vez más grande y más grande, contamos con que en
- Nuestro ejemplo el cociente sea cada vez más cerca de 1/2.

Variables aleatorias
Probabilidad

- Esto nos permite " definir " la probabilidad Pr(H) de obtener de un tiró “cabezas” como el límite, cuando N tiende al infinito, de esta secuencia de cocientes:

$$\Pr(H) = \lim_{N \to \infty} \frac{N_H}{N}$$

•En la práctica actual, por supuesto, no podemos tirar una moneda un número infinito de veces; por lo tanto en términos generales, esta fórmula se aplica lo más exactamente posible a las situaciones en las cuales hemos asignado ya una probabilidad a priori a un resultado particular (en este caso, nuestra suposición de que la moneda era una moneda " justa").

Variables aleatorias
Probabilidad

- La ley de números grandes entonces dice esto, dado Pr(H), y cualquier número arbitrariamente pequeño ε, existe un cierto número *n* tal que para todo el *N* > *n*

$$\left|\Pr(H) - \frac{N_H}{N}\right| < \epsilon$$

- En otras palabras proclamando " la probabilidad de cabezas es 1/2 ", nosotros afirmamos que, si tiramos nuestra moneda muchas veces, el número de “cabezas” sobre el número de tiros totales se irá acercando arbitrariamente a 1/2; y entonces permanecerá *por lo menos* cerca de 1/2 mientras guardamos el realizar tiros adicionales de la moneda.

Variables aleatorias

La definición moderna

- Al principio se debe definir un espacio Ω para el conjunto de todos los acontecimientos elementales posibles X_i tales que sean exclusivos, esto es, que la ocurrencia de uno de ellos implica que los demás no suceden. Entonces definimos la probabilidad de la ocurrencia de X_i, $P(X_i)$, con las siguientes características:

- $P(X_i) >= 0$ para todos i
- $P(X_i \text{ or } X_j) = P(X_i) + P(X_j)$ (1.1)
- $\Sigma P(X_i) = 1$.

Variables aleatorias Continuas

- Para las variables aleatorias continuas (que intervalos continuos cubran los posible valores) necesitamos las herramientas *de la función de densidad de probabilidad* y su integral, *la función de distribución acumulativa*
- **Función de la densidad de la probabilidad**
- $f(x_0) = P(x_0 - \Delta x < x < x_0 + \Delta x)/\Delta x$

Variables aleatorias Continuas

- **Distribuciones acumulativas**
- La variable aleatoria X es caracterizada o por su pdf f(X) o su distribución acumulativa F(X). Donde:
- $F(X) = \int f(X)dX.$ (2,16)
- Por la construcción
- $F(X_{min}) = 0$, $F(X_{max}) = 1$
- Si la gama de valores posibles es $X_{min} < X < X_{max}$. F(X) es una función monótona de X tal que
- $F(X_1) > F(X)$ para todos $X_1 > X$
- De la definición es obvio que $F(X_1)$ es la probabilidad de X sea más pequeño que X_1

Características de las distribuciones

- **Esperanza, promedio y varianza**
- Las funciones de la densidad de la probabilidad se utilizan como funciones generadoras para obtener la información sobre las variables al aleatorias. Si g(X) es una cierta función de una variable aleatoria X con la densidad f(X), la esperanza de g(X) es el número
- $E(g) = \int g(X)f(X)dX,$ (2,22)
- Donde es la integración sobre el todo el espacio de X.
- La esperanza E es un operador linear
- $E[ag(X) + bh(X)] = aE[g(X)] + bE[h(X)]$ (2,23)
- La esperanza de la variable aleatoria X por sí mismo se llama el promedio de la densidad f(X) o el valor previsto de X para la densidad f(X), y es denotada por el μ.

Características de distribuciones

- $\mu = \int Xf(X)dX$.
- La esperanza de la función $(X - \mu)^2$ se llama la varianza V(X) de la densidad f(X)
- $V(X) = \sigma^2 = E[(X - \mu)^2] = E[X^2 - 2X\mu + \mu^2] = E[X^2] - \mu^2 = \int(X-\mu)^2 f(X)dX$
- La cantidad σ se llama *desviación estándar.*
- Observe que el promedio de una muestra de la distribución no siempre puede existir. Un ejemplo es dado por la densidad de la distribución de Cauchy
- $f(X) = 1/p(1 + X^2)$
- Cabe destacar que f(X) es la fórmula de Breit-Wigner encontrada a menudo en la física para describir amplitudes de la interacción de las partículas de la resonancia.

El teorema de limite central

- Este teorema es de gran alcance de la importancia central en problemas teóricos y prácticos en estadística. Si tenemos una secuencia de variables aleatorias independientes X_i, cada uno con una distribución con valor medio μ_i y varianza $\sigma^2{}_i$, entonces la distribución de la suma $S = \Sigma$ Xi tendrá un valor medio $\Sigma\mu_i$ y una varianza $\Sigma\sigma^2{}_i$. El teorema de limite central indica como la suma se distribuye en el limite con N grande. Es decir N —> $\propto$
-
- $(S - \Sigma X_i)/\sqrt{(\Sigma\sigma^2{}_i)}$ —> N(0,1)

- No daré la prueba que se basa en el uso de funciones características.

Variables aleatorias Discretas

Función lineal de variables aleatorias

La expectativa de la función lineal de varias variables aleatorias $X_1, \ldots, X_N$

$E[\Sigma a_i X_i] = \Sigma a_i E[X_i] = \Sigma a_i \mu_{Xi}$

La varianza es

$V(\Sigma a_i Xi) = \Sigma a_i^2 V[Xi] + 2 \Sigma \Sigma a_i a_j \operatorname{cov}(Xi,Xj)$

Si las Xi son independientes sin correlación la ecuación de arriba toma la simple forma

$V(\Sigma a_i Xi) = \Sigma a_i^2 V[Xi]$

En el caso cuando las Xi son diversas pruebas N del mismo experimento, el promedio que es dado por el valor medio de las observaciones Xi,

Variables aleatorias Discretos

$E[Xi] = \mu$ y varianza $V(Xi) = \sigma^2$ para todo i

La expectativa del promedio de la observación de N entonces es

$E[\Sigma Xi/N] = \Sigma E[Xi]/N = N\mu/N = \mu$

y la varianza del promedio

$V(\Sigma Xi/N) = \Sigma V[Xi]/N^2 + 2/N^2 \ \Sigma\Sigma cov(Xi,Xj) = \sigma^2/N + 2/N^2 \ \Sigma\Sigma cov(Xi,Xj)$

Si los ensayos son independientes, cov(Xi,Xj) = 0 para cada par (i,j), y el término doble de la suma cae hacia fuera. Entonces tenemos el resultado bien conocido.

$\sigma(\Sigma Xi/N) = \sigma(Xi)/\sqrt{N}$

Lo cuál dice que la desviación de estándar del medio disminuye con $\sqrt{N}$ mientras que N aumenta.

Distribuciones extensamente usadas

- **Distribución binomial**
- Función de la probabilidad para la distribución binomial:
-
- $P(r) = B\{N,r\}p^r(1\text{-}p)^{N\text{-}r}$ donde $B\{N,r\} = N!/r!(N\text{-}r)!$, $r = 0,1,2,$ ……,N
- $0 < p < 1$
- Esperanza matematica E[r] = Np
- Varianza V(r) = Np(1-P)

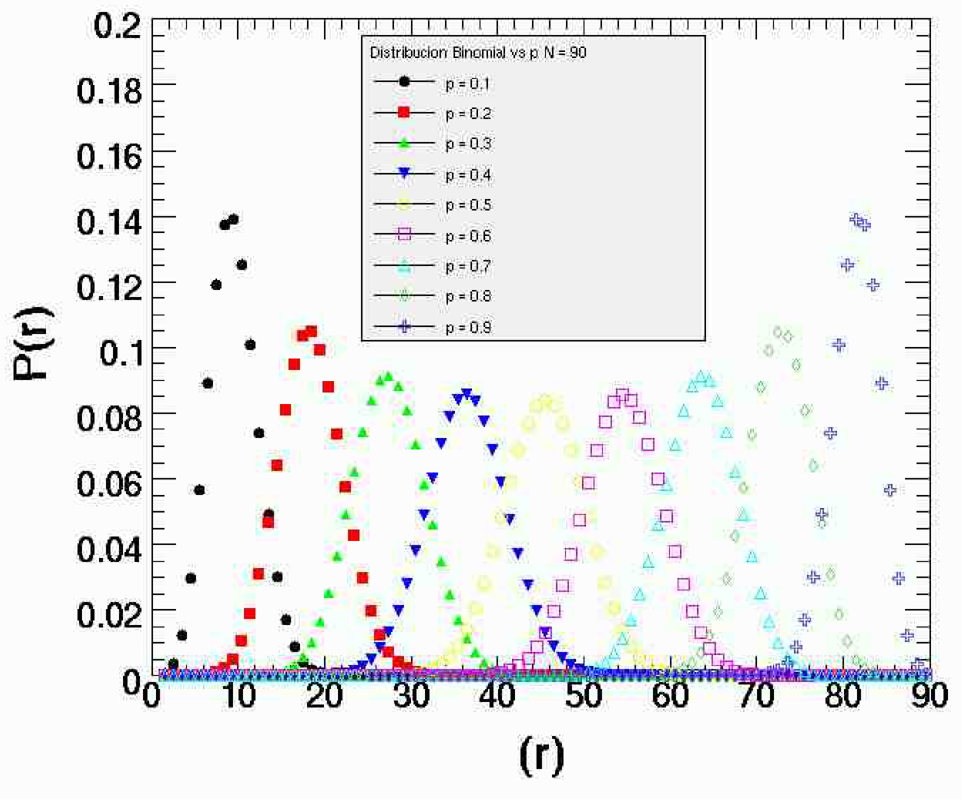
Distribuciones extensamente usadas

- La distribución binomial de la probabilidad de encontrar exactamente r éxitos en N ensayos, cuando la probabilidad del éxito en cada solo ensayo es una constante, p. la distribución del número de acontecimientos en un solo compartimiento del histograma es binomial.
- Si p es desconocido, una estimación imparcial de la varianza se da por
- $V(r) = N^2 (r/N) (1-r/N) / (N-1)$
- Considere un estudio de la absorción de la radiografía en una capa. Suponga que la probabilidad de la absorción de un solo fotón es p. el número de fotones que atraviesan la capa se distribuye según la ley binomial:

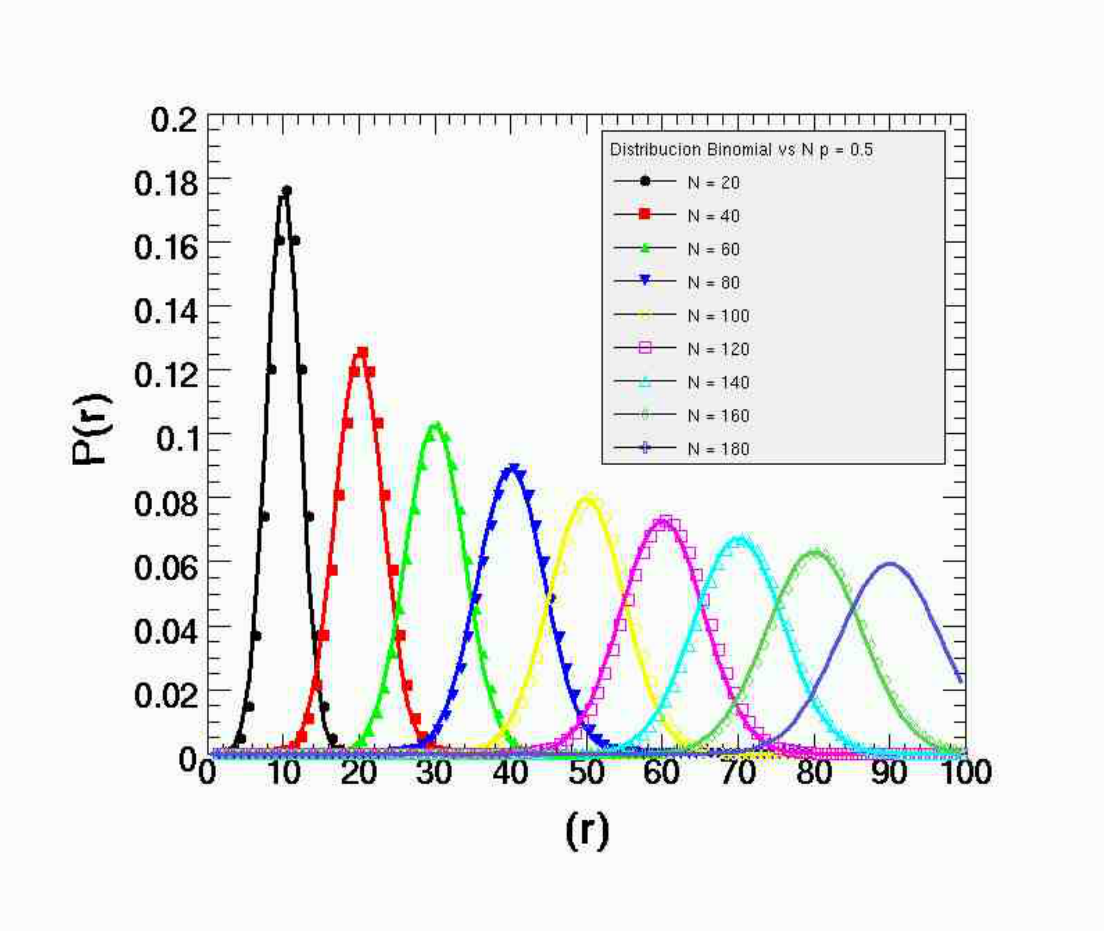
Distribuciones extensamente usadas

- $P(N) = \{N_0\ N\}\ (1\text{-}p)^N\ (p)^{N_0 - N}$
- Entonces el esperanza matemática para N es $N_0(1\text{-}p)$, donde N_0 es el número de fotones antes de la capa. De la varianza la esperanza Np, la desviación estándar está dada por $\sqrt{Np}$ (no por $\sqrt{N}$)
- Entonces para las capas gruesas cuando $p \approx 1$, el valor $\sqrt{N}$ para la desviación estándar es apropiado!

Distribuciones extensamente usadas



Distribuciones extensamente usadas



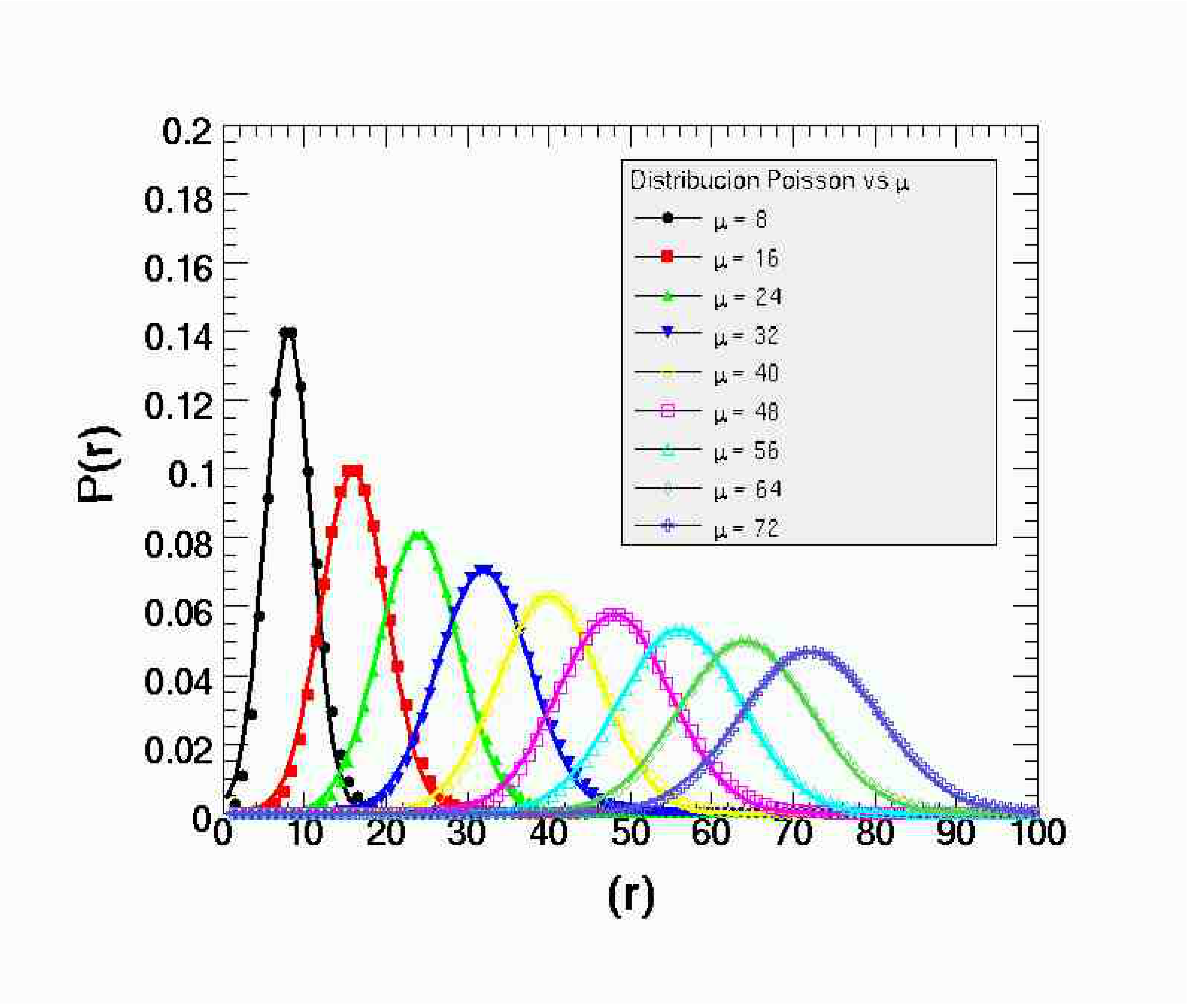
Distribuciones extensamente usadas

- **Distribución de Poisson**
- Función de probabilidad
- $P(r) = \exp(-\mu)\ r^{-\mu} / r!$
- Esperanza matematica $E[r] = \mu$
- Varianza $V(r) = \mu$
- La distribución de Poisson da la probabilidad de encontrar exactamente r acontecimientos en una longitud del tiempo dada, si ocurren los acontecimientos independientemente, a una tasa constante.

Distribuciones extensamente usadas

- Es un caso límite de la distribución binomial para p - —> 0 y N —> $\propto$ cuando Np = μ una constante finita.
- Cuando μ —> $\propto$ la distribución de Poisson converge a la distribución normal.
- Ejemplo:
- Suponga que son las partículas emitas de una fuente radiactiva en un índice medio de g partículas por unidad de tiempo, de una manera tal que la probabilidad de la emisión en δt sea el $\gamma\delta t$, y la probabilidad de más de una emisión en despegue es $O(\delta t^2)$. Entonces la distribución del número de partículas X, emitidas en un intervalo fijo t del tiempo, es Poisson, con g malo t:
- $P(X = r) = (\gamma t)^r \exp(-\gamma t)/r!$

Distribuciones extensamente usadas



Distribuciones extensamente usadas

- **Normal uno dimensional**
- Función de densidad de la probabilidad para la distribución (Gaussiana) normal:
- $f(X) = N(\mu,\sigma^2) = \exp(-(X-\mu)^2/2\sigma^2)/\sigma\sqrt{2\pi}$ donde μ - número real, σ - número real positivo
- Distribución acumulativa
- $\Phi(X) = \Phi((X-\mu)/\sigma)$ donde $\Phi(Z) = \int \exp(-x^2/2)dx/\sqrt{2\pi}$ donde integración - del $\propto$ a Z
- Esperanza matematica $E[X] = \mu$
- Varianza $V(X) = \sigma^2$

Distribuciones extensamente usadas

- La distribución teórica más importante de la estadística es la función normal de la densidad de la probabilidad, o Gaussiana, generalmente abreviado $N(\mu, \sigma^2)$. Su distribución acumulativa, se llama *la probabilidad normal integral* o *la función de error.*
- La desviación de estándar σ no es la anchura del p.d.f. en la mitad de la altura. La anchura en la mitad de la altura es $1{,}76\sigma$ El contenido de la probabilidad de varios intervalos se da abajo:

Distribuciones extensamente usadas

- P(-1.64 ≤ (x m) / s) ≤ 1,64) = 0,900
- P(-1.96 ≤ (x m) / s) ≤ 1,96) = 0,950
- P(-2.58 ≤ (x m) / s) ≤ 2,58) = 0,990
- P(-3.29 ≤ (x m) / s) ≤ 3,29) = 0,999

- La función N(0,1) se llama la densidad normal estándar y su función acumulativa,
- $\Phi(X) = \int \exp(-t^2/2)dt/\sqrt{2\pi}$ donde integración - de $\propto$ a X
- Se llama la distribución normal estándar.
- Cualquier combinación linear del Xi es también normal. Si Z = $a_1X_1 + a_2X_2$ entonces Z es también normal con valor medio $a_1\mu_1 + a_2\mu_2$ y varianza $a^2_1\sigma^2_1 + a^2_2\sigma^2_2$.

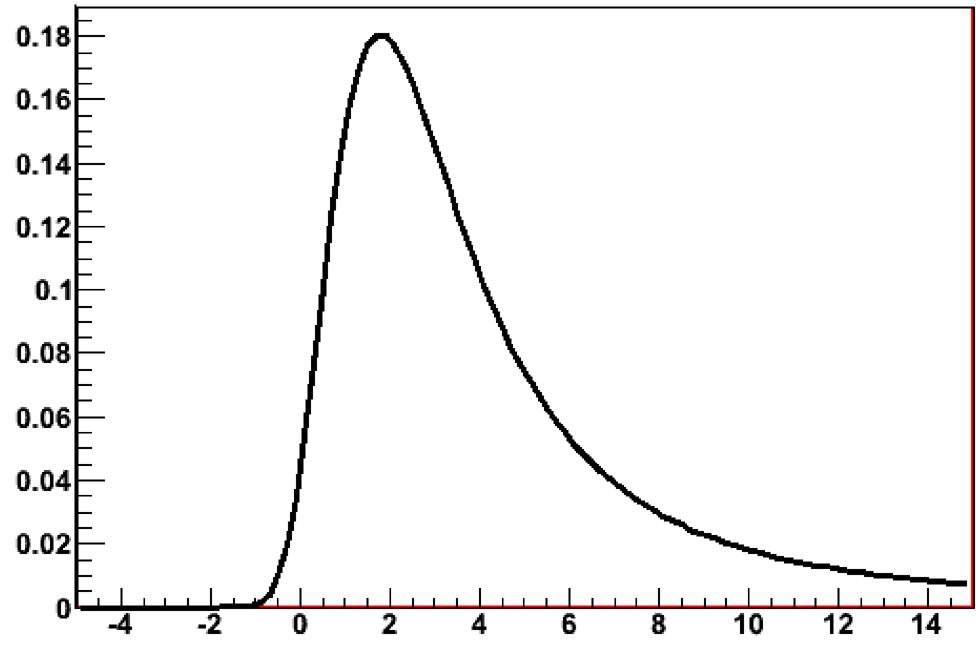
Distribuciones extensamente usadas

- Distribución de Landau
- Función de la densidad de la probabilidad:

$$p(x) = \frac{1}{\pi} \int_0^\infty e^{-t \log t - xt} \sin(\pi t)\, dt.$$

$$\varphi(t;\mu,c) = \exp\Big[\, it\mu - |c\,t|(1 + \tfrac{2i}{\pi}\log(|t|)\Big].$$

M = 2 y c = 1



Distribuciones extensamente usadas

- Distribución de Landau
- Distribución de la deposición de energía por piones de 500 MeV en detectores de Si de diferentes espesores

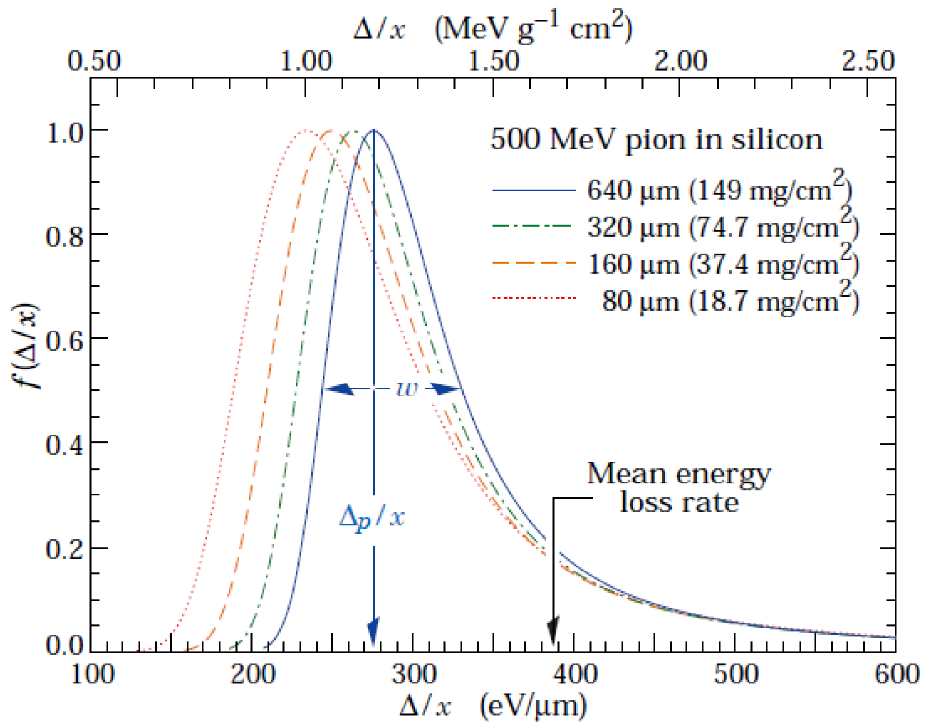


Figure 27.7: Straggling functions in silicon for 500 MeV pions, normalized to unity at the most probable value δ_p/x. The width w is the full width at half maximum.

Distribuciones extensamente usadas

- **Distribución del Chi-cuadrado**
- Función de la densidad de la probabilidad:
- $f(X) = (X/2)^{N/2-1}\exp(-X/2)/2\Gamma(N/2)$ donde número – verdadero positivo de X, N - número entero positivo (grados de la libertad)

- Esperanza Matematica E[X] = N
- Varianza V(x) = 2N
- Suponga ese $X_1,\ldots., X_N$ es las variables normales independientes, estándares, N(0,1). Entonces la suma de cuadrados

Distribuciones extensamente usadas

- $X^2_{(N)} = \Sigma\, Xi^2$
- Es triste tener una distribución del chi-cuadrado $\chi^2(N)$, con N grados de libertad. Si $X^2_{(N)}$ y $X^2_{(M)}$ tienen distribuciones independientes de χ^2 con N y M grados de libertad, respectivamente, entonces la suma
- $X^2_{(K)} = X^2_{(N)} + X^2_{(M)}$
- Tiene una distribución de χ^2 con K=N+M grados de libertad.
- Para N grande la cantidad
- $Z_N = (X^2(N) - N)/\sqrt{2N}$
- Es la norma estándar N(0,1)

Distribuciones extensamente usadas