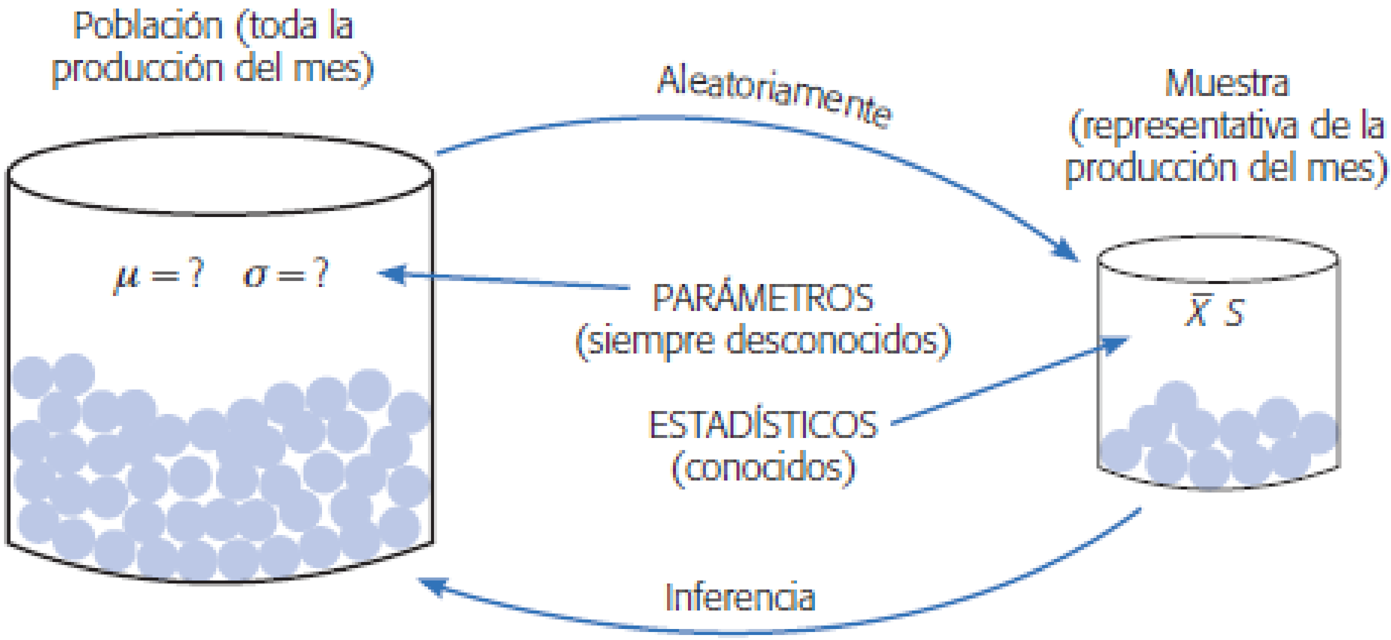
Introducción al Diseño de Experimentos

www.academia.utp.ac.pa/humberto-alvarez

Introducción

- Una *población* o *universo* es una colección o totalidad de posibles individuos, especímenes, objetos o medidas de interés sobre los que se hace un estudio.
- Las poblaciones pueden ser:
  - *Finita*: pueden medir todos los individuos para tener un conocimiento "exacto" de sus características (*parámetros*).
  - *infinita* se tendrá que sacar una *muestra representativa* de dicha población, y con base en las características medidas en la muestra (*estadísticos*) se podrán hacer afirmaciones acerca de los parámetros de la población. Un *estadístico* se define como cualquier función de los datos muestrales que no contiene parámetros desconocidos.
- La *inferencia estadística* hace afirmaciones válidas acerca de la población o proceso con base en la información contenida en una muestra o representación válida de la población.

Relación entre población y muestra



Distribuciones de probabilidad

- Una variable aleatoria es una función que asocia un número real con cada elemento del espacio muestral de un experimento
- La *distribución de probabilidad* o *distribución de una variable aleatoria X* relaciona el conjunto de valores posibles de *X* (rango de *X*), con la probabilidad asociada a cada uno de estos valores y los representa a través de una tabla o por medio de una función planteada como una fórmula.
  - **Variable aleatoria discreta:** si se puede contar su conjunto de resultados posible o si están relacionadas con el conjunto de números enteros.
  - **Variable aleatoria continua**: Cuando puede tomar valores en la escala continua o está relacionada con un conjunto continuo de valores.

Distribuciones de probabilidad

- Una variable aleatoria toma cada uno de sus valores con cierta probabilidad.
- La función de probabilidad es la representación de todas las probabilidades de una variable aleatoria X mediante una expresión matemática tal que

$$P(X = x) = f(x)$$

Distribuciones de probabilidad discretas

- El conjunto de pares ordenados $(x, f(x))$ es una función de probabilidad discreta o distribución de probabilidad discreta de la variable discreta X, si para cada resultado x
  1. $f(x) \geq 0$
  2. $\Sigma_x f(x) = 1$
  3. $P(X = x) = f(x)$
- La distribución acumulada F(x) de una variable aleatoria discreta X con distribución de probabilidad $f(x)$ es:

$$F(x) = P(X \leq x) = \Sigma_{t \leq x} f(t) \quad \text{para } -\infty \leq x \leq \infty$$

Media de una variable aleatoria

- La media de una variable aleatoria es el valor más esperado de dicha variable.

$$\mu = E(X) = \sum_{x} x f(x)$$

Varianza de una variable aleatoria

- Sea X una variable aleatoria con distribución de probabilidad f(x) y media μ, la varianza de X será:

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 f(x)$$

Donde la desviación estándar σ será la raíz cuadrada de la varianza

Distribución uniforme discreta

- Si la variable aleatoria discreta X toma valores discretos x_1, x_2, ….., x_k con igual probabilidad, entonces estos valores están distribuidos en función a la Distribución Uniforme Discreta:

$$P(X = x) = f(x;k) = \frac{1}{k}$$

$$\mu = \frac{\sum_{i=1}^{k} x_i}{k}$$

$$\sigma^2 = \frac{\sum_{i=1}^{k} (x_i - \mu)^2}{k}$$

El Proceso de Bernoulli

- El experimento consiste en n pruebas iguales que se repiten
- Cada prueba produce un resultado que se puede clasificar como éxito o fracaso
- La probabilidad de un éxito, p, permanece constante en cada prueba
- Cada prueba es independiente

La Distribución Binomial

- En experimento de Bernoulli puede tener como resultado un éxito con probabilidad p y un fracaso con probabilidad 1-p. Entonces la distribución de probabilidad de la variable aleatoria X, el número de éxitos en n pruebas independientes es:

$$P(X = x) = f(x; n, x) = \binom{n}{x} p^x q^{n-x}$$

donde

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

$$\mu = np$$

$$\sigma^2 = npq$$

$$P(x \leq x) = \sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k}$$

Proceso de Poisson

- Son aquellos experimentos donde los resultados ocurren durante un intervalo dado o en una región específica.
- Los resultados que ocurren en un intervalo son independientes de los resultados en otro intervalo o región.
- La probabilidad de que ocurra un solo resultado durante un intervalo muy corto es proporcional a la longitud del intervalo y no depende del número de resultados en otros intervalos
- La probabilidad de que se den resultados simultáneos en un intervalo es despreciable.

Distribución de Poisson

- Sea X la variable aleatoria asociada con la ocurrencia de eventos en un proceso de Poisson. Esta variable estará distribuida de acuerdo a la siguiente función:

$$P(X=x)=\frac{e^{-1/\lambda}(1/\lambda)^{x}}{x!}, \text{para } x=1,2,3,\ldots$$

$$P(X\leq x)=\sum_{k=0}^{x}\frac{e^{-1/\lambda}(1/\lambda)^{k}}{k!}$$

$$\mu=\sigma\ \ =1/\lambda$$

Donde $1/\lambda$ es el número promedio de eventos por unidad de tiempo

Distribuciones de probabilidad continuas

- El conjunto de pares ordenados $(x, f(x))$ es una función de probabilidad continua o función de densidad de probabilidad de la variable aleatoria continua X, definida en el conjunto de números reales $\Re$, si:
  1. $f(x) \geq 0$, para toda $x \in \Re$
  2. $\int_{\Re} f(x) = 1$
  3. $P(a < x < b) = \int_{a < x < b} f(x)dx$
- La distribución acumulada F(x) de una variable aleatoria continua X con distribución de probabilidad $f(x)$ es:

$$F(x) = P(X \leq x) = \int_{t \leq x} f(t)dt \quad \text{para } -\infty \leq x \leq \infty$$

Media de una variable aleatoria

- La media de una variable aleatoria es el valor más esperado de dicha variable.

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Varianza de una variable aleatoria

- Sea X una variable aleatoria con distribución de probabilidad f(x) y media μ, la varianza de X será:

$$\sigma^2 = E\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

Donde la desviación estándar σ será la raíz cuadrada de la varianza

Distribución uniforme

- Una variable aleatoria continua X tendrá una distribución uniforme en el intervalo [a, b] de la forma:

$$f(x) = \frac{1}{b-a}$$

$$P(X \leq x) = \int_0^x \frac{dy}{b-a}$$

$$\mu = \frac{a+b}{2}, \sigma^2 = \frac{(b-a)^2}{12}$$

Distribución exponencial

- La variable aleatoria continua X tiene una **distribución exponencial** con parámetro λ si su función de probabilidad está dada por:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

$$P(X \le x) = \int_0^x \lambda e^{-\frac{1}{\lambda} y} dy$$

$$P(0 \le X \le x) = 1 - e^{-\frac{1}{\lambda} x}$$

$$\mu = \frac{1}{\lambda}, \ \sigma = \frac{1}{\lambda}$$

Donde 1/λ es el número promedio de eventos por unidad de tiempo

Distribución normal

- Es la distribución más importante en el campo de la estadística
- La curva normal describe aproximadamente muchos fenómenos que ocurren en la naturaleza
- Desarrollada en 1733 por Abraham DeMoivre
- Aunque fue Karl Fiedrich Gauss quien demostró su aplicabilidad

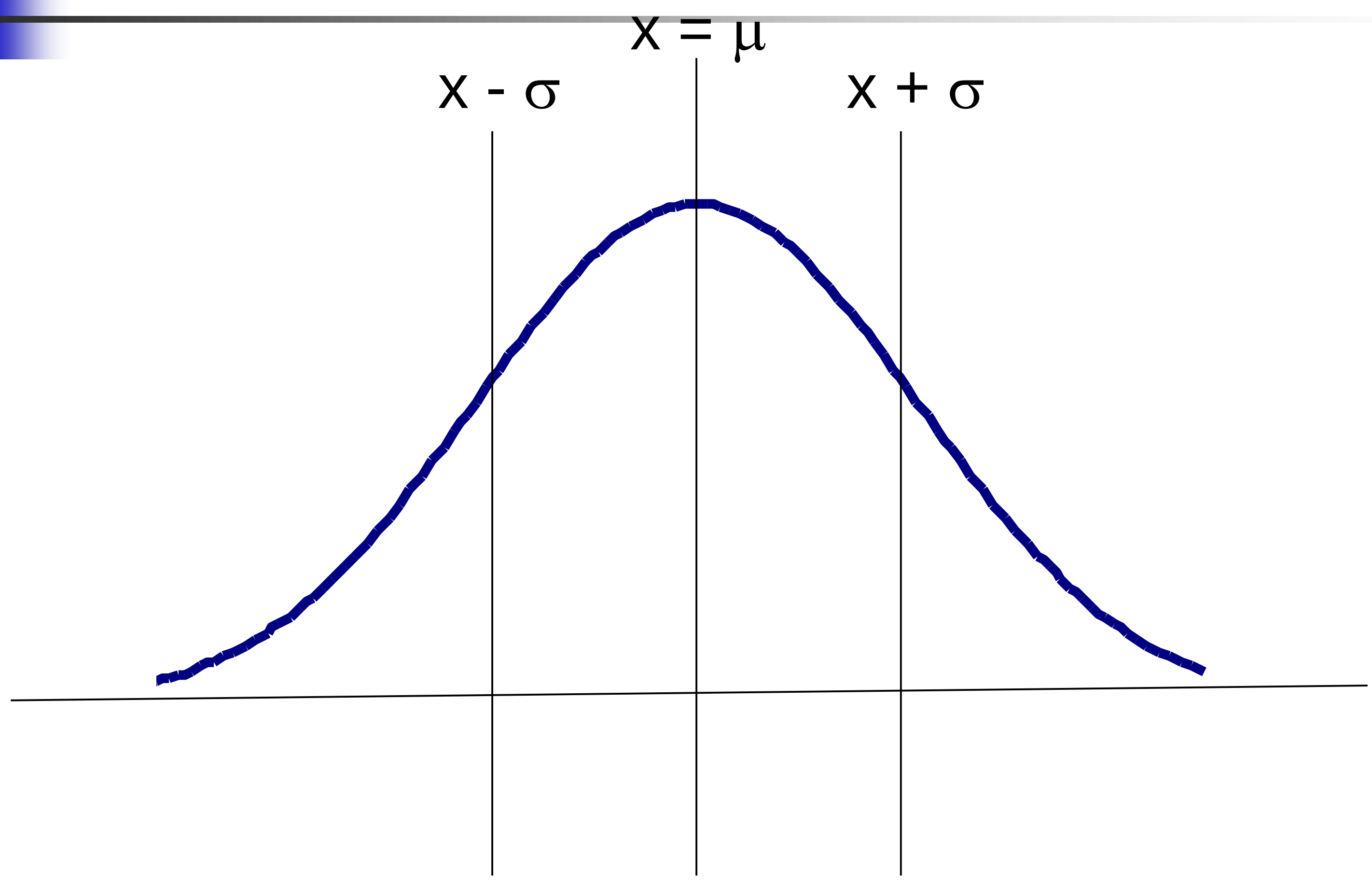
Distribución normal

- La función de probabilidad de la variable aleatoria normal X, con media μ y varianza σ^2 está dada por:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}, \quad -\infty < x < \infty$$

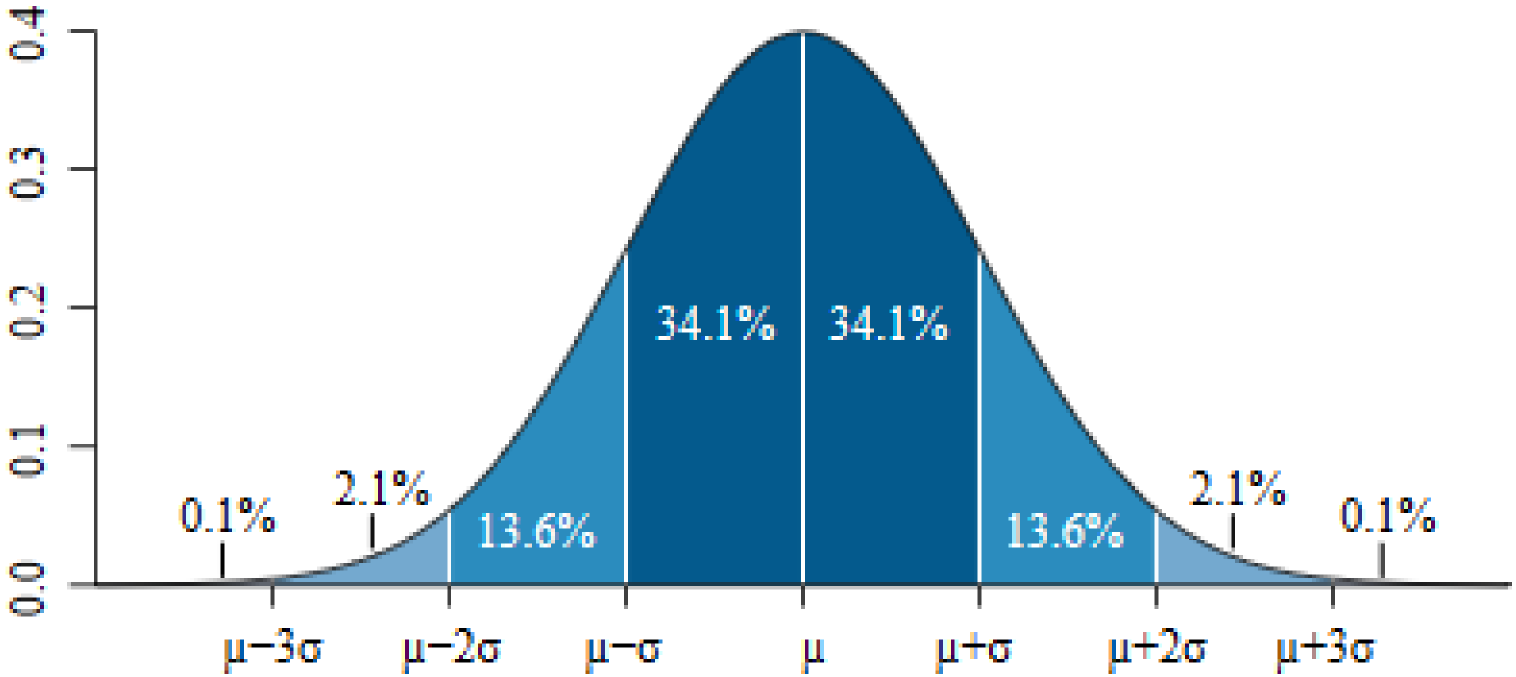
Características de la curva normal

- La variable aleatoria normal X se expresa como $N(\mu, \sigma^2)$
- La moda está localizada en el punto donde $x = \mu$
- La curva es simétrica en el eje de la media μ
- Sus puntos de inflexión están en $x = \mu \pm \sigma$
- Se aproxima al eje horizontal de manera asintótica en ambos lados
- Su área total es igual a 1

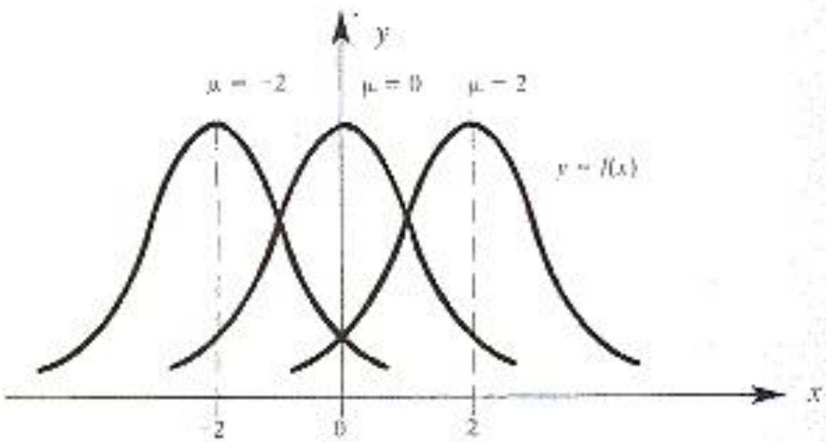


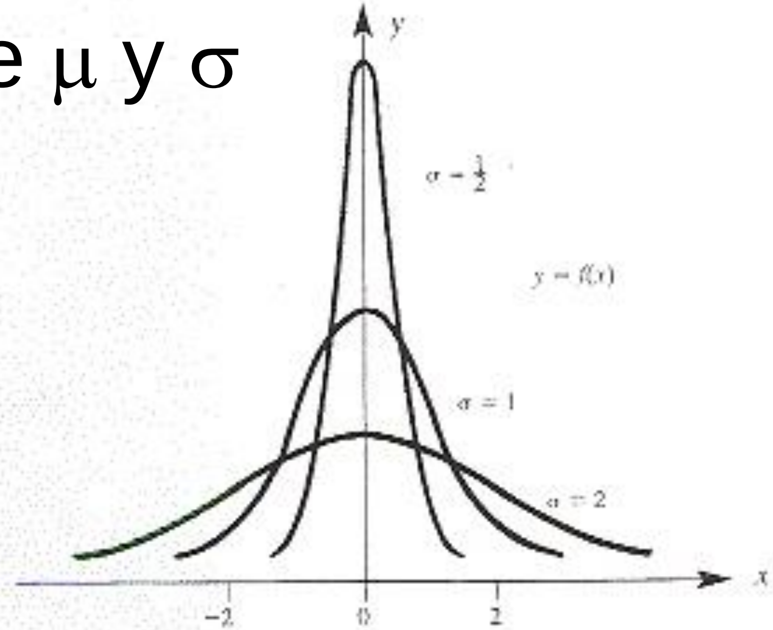
X = μ
X - σ
X + σ

Probabilidades alrededor del eje central o media



Efectos de μ y σ





Áreas bajo la curva normal

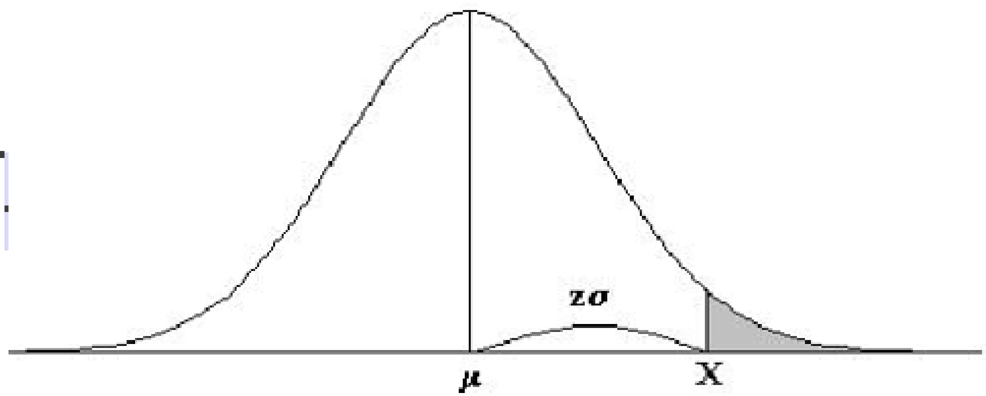
- Es matemáticamente difícil calcular las probabilidades o el área bajo la curva normal:

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi\sigma}} \int_{x_1}^{x_2} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx$$

- Se determina el valor normalizado estándar tal que N(0,1)

$$Z = \frac{x-\mu}{\sigma}$$

Áreas bajo la curva normal



Ejemplo:

$$Z = \frac{X - \mu}{\sigma}$$

P [Z > 1] = 0.1587
P [Z > 1.96] = 0.0250

Desv. normal x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681

Desv. normal x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010

=distr.norm.inv(p, μ, σ)