

# Modelos No Lineales

## Clase Modelo

**Edinson Tolentino**

email: [edinson.tolentino@gmail.com](mailto:edinson.tolentino@gmail.com)

Twitter: [@edutoleraymondi](https://twitter.com/edutoleraymondi)

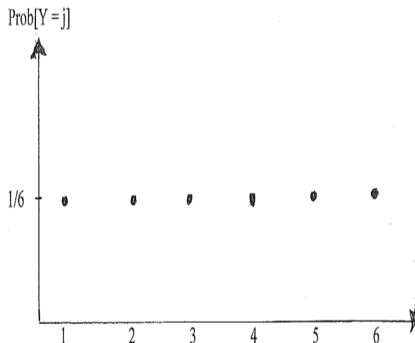
Educate Peru

- 1 Conceptos Básicos estadísticos
- 2 Reintroducción Maximun Likelihood (ML)
- 3 Modelo de Probabilidad Lineal
- 4 Modelo de Regresión Logística
- 5 Modelo Probit
- 6 Probabilidades y características promedio
- 7 Efectos Marginales

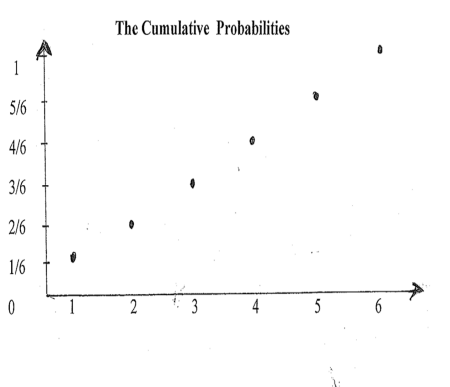
- Asuma el lanzamiento de un dado , el cual denota 6 lados como producto
- La formula para la probabilidad de cada producto estará dado por:

$$Prob[Y = j] = \frac{1}{6}$$

- Donde,  $j = 1, 2, 3, 4, 5, 6$



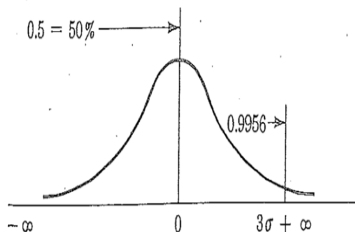
- Probabilidad acumulada



- La función densidad de probabilidad (pdf) y el gráfico de la función de probabilidad acumulada (CDF) de una distribución normal.

$$pdf : f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) \Rightarrow z \sim N(0, 1)$$

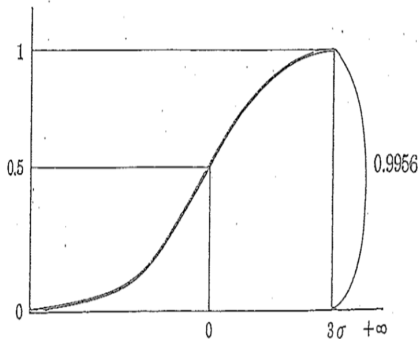
The standard normal probability density curve



- La función de probabilidad acumulada (CDF) de una distribución normal.

$$CDF : F(z) = \int_{-\infty}^z f(z) dz$$

- La CDF de la distribución normal estandar tiene una forma de una curva de cambio



- Maximum Likelihood (ML) requiere un conocimiento exacto de la función de distribución de probabilidad de una variable subyacente.
- La pdf puede ser representado por una curva

$$prob(a < Y < b) = \int_a^b f(y)dy$$

$$prob(a < Y < a) = \int_a^a f(y)dy = 0$$

$$prob(-\infty < Y < +\infty) = \int_{-\infty}^{+\infty} f(y)dy = 1$$

- Supuesto de la distribución dentro del rango  $-\infty$  hacia  $+\infty$

- Entonces:

$$prob(Y \leq c) = \int_{-\infty}^c f(y) dy = prob(Y < c)$$

$$F(c) = \int_{-\infty}^c f(y) dy$$

- Donde  $F(c)$  represente la probabilidad acumulada
- $F(c)$  es llamada la función de probabilidad acumulada (operador) , llamada CDF
- Esto permite que, la derivada de la CDF es el pdf:

$$\frac{\partial F(y)}{\partial y} = f(y)$$



- **Ejemplo:** Asuma que  $Y$  es una variable aleatoria con valor de 0 o 1, donde 1 ocurre dado una probabilidad de  $\lambda$  y 0 con probabilidad de  $1 - \lambda$

- **Ejemplo:** Asuma que  $Y$  es una variable aleatoria con valor de 0 o 1, donde 1 ocurre dado una probabilidad de  $\lambda$  y 0 con probabilidad de  $1 - \lambda$
- La distribución de la variable aleatoria  $Y$  es caracterizada por un parametro  $\lambda$ , el cual significa

$$Prob[Y = 1] = \lambda$$

$$Prob[Y = 0] = 1 - \lambda$$

- Asumiendo que se tiene tres valores de  $Y$ , de forma aleatoria, los cuales son:

$$Y = (1, 1, 0)$$

- ¿Cuál es el estimador de Maximum likelihood de  $\lambda$ ?

- **Paso 1:** Denotar la función de verosimilitud (*Likelihood Function*), probabilidad conjunta de las observaciones sera :
- Entonces:

$$LF = \text{Prob}(Y = 1) \times \text{Prob}(Y = 1) \times \text{Prob}(Y = 0)$$

- Luego:

$$LF = \lambda \times \lambda \times (1 - \lambda)$$

- Por lo tanto:

$$LF = \lambda^2(1 - \lambda)$$

- **Paso 2:** optimizar la función de verosimilitud (*Likelihood Function*):
- Por lo tanto:

$$LF = \lambda^2(1 - \lambda) = \lambda^2 - \lambda^3$$

- Optimización:

$$\frac{\partial LF}{\partial \lambda} = 2\lambda - 3\lambda^2 = 0$$

- Por tanto, se encuentra el valor de  $\lambda$  que hace óptimo  $LF$

- La metodología de ML requiere tres puntos:
  - 1 Se debe especificar la pdf para casos continuos (o la CDF para casos discretos) dada una función donde no se conoce los parametros  $\beta_1, \dots, \beta_k$
  - 2 El uso de la diferenciación (derivadas) para encontrar el valor de los parametros que optimizan (maximizan) la función de verosimilitud ( $FL$ )
  - 3 Uso de las ecuaciones de las derivadas parciales para encontrar los demas parametros
  - 4 Esto requiere de iteraciones no lineales , por tanto , por lo general se usa un  $\text{Log}\vartheta$  para para función de verosimilitud ( $L$ )

- Entonces, si se tiene el parametro  $\beta_0$  como estimador de  $ML$  y  $L$  como el logaritmo de la *función verosimilitud* (Likelihood function,  $LF$ ), los elementos que requiere el método de  $ML$  son:

- La **función de score** dado por:

$$\frac{\partial L}{\partial \beta_0} = S(\beta_0)$$

- La **información** dado por :

$$-E \frac{\partial^2 L}{\partial \beta_0^2} = I(\beta_0)$$

**Nota:** (Conocido como la información de Fisher)

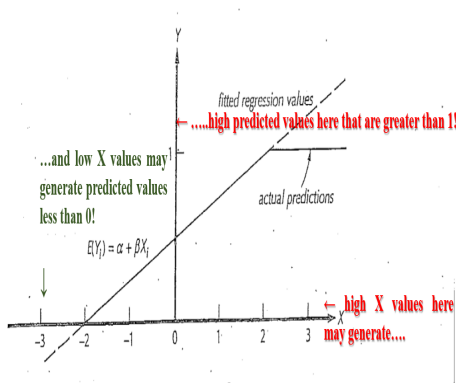
- La **varianza muestral asintótica** para los estimadores de  $ML$  obtenidos como;

$$Var(\beta_0) = [I(\beta_0)]^{-1}$$

- Siguiendo la especificación:

$$y_i = X_i' \beta + \mu_i$$

- Donde :  $y = 1$  o  $y = 0$  ,  $E(\mu) = 0$ 
  - ▶ La probabilidad de que  $y = 1$  es  $\rho$
  - ▶ Mientras que la probabilidad de que  $y = 0$  sera  $1 - \rho$
- Entonces :  $E[y_i] = X_i' \beta = \text{Prob}(y = 1) = \rho_i$



- **Caso 1:** En un encuentro de fútbol entre el Arsenal y el Fullham, las casas de apuestas ofrecen una probabilidad (odds) de **2/9** en contra de que el Arsenal gane (local)
  - ▶ El enunciado anterior significa que si tu apuestas un dolar 1 \$ sobre una victoria del Arsenal, las casas de apuestas te pagaran 1 \$ mas 0.22 centavos
  - ▶ Por tanto, las casas de apuestas tienen un **odds ratio** en contra de que ocurra el evento (victoria de arsenal) definido como:

$$\frac{(1 - \rho)}{\rho} = \frac{2}{9}$$

- ▶ La probabilidad implícita de la victoria del local (Arsenal) que las casas de apuestas usan es:

$$\rho = \frac{9}{11} = 0.818$$

- ▶ las casas de apuestas asumen una victoria del Arsenal en 82 %



- La curva de CDF para un modelo de regresión logística está dado por:

$$Prob(y_i = 1) = F(X' \beta) = \frac{\exp(X' \beta)}{1 + \exp(X' \beta)}$$

**Nota** Probabilidad para cuando el evento ocurra

- Por otro lado:

$$\begin{aligned} Prob(y_i = 0) &= 1 - F(X' \beta) = \\ &= 1 - \frac{\exp(X' \beta)}{1 + \exp(X' \beta)} \\ &= \frac{1 + \exp(X' \beta)}{1 + \exp(X' \beta)} - \frac{\exp(X' \beta)}{1 + \exp(X' \beta)} \\ &= \frac{1}{1 + \exp(X' \beta)} \end{aligned}$$

**Nota** Probabilidad para cuando el evento no ocurra

- Usando el modelo logístico, podemos construir el siguiente ratio:

$$\frac{Prob(y_i = 1)}{1 - Prob(y_i = 1)} = \exp(X' \beta)$$

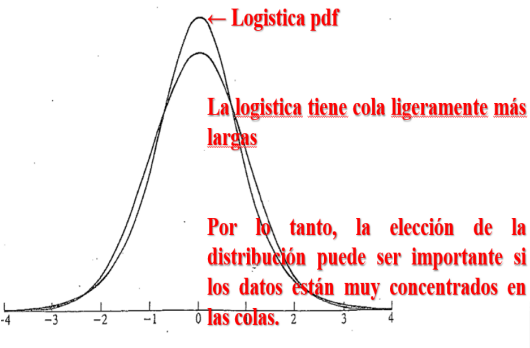
- Luego podemos tomar el logaritmo a la ultima expresión:

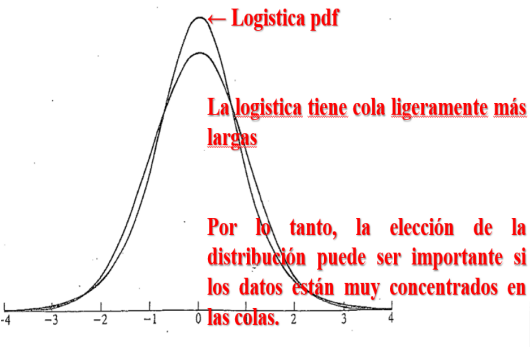
$$\log \left( \frac{Prob(y_i = 1)}{1 - Prob(y_i = 1)} \right) = X' \beta$$

- Si tomamos diferencial respecto a la variable explicativa , se tiene:

$$\frac{\partial \log \left( \frac{Prob(y_i=1)}{1-Prob(y_i=1)} \right)}{\partial X'} = \beta$$

Modelo	Interpretación de coeficientes
Modelo de Probabilidad Lineal	El efecto de un pequeño cambio en el regresor sobre <b>la probabilidad del evento de ocurrencia</b>
Modelo logit	El efecto de un pequeño cambio en el regresor sobre <b>el log odds ratio del evento de ocurrencia</b>



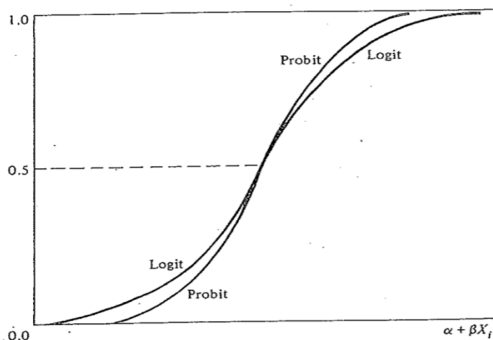


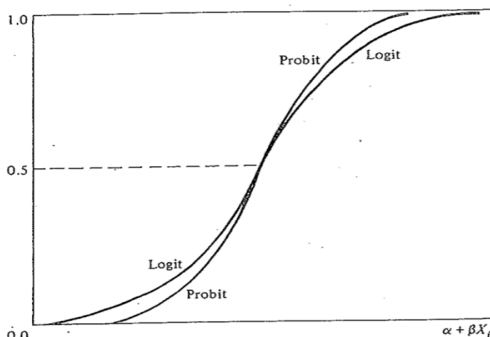
- Probit

$$\sim N(0, 1)$$

- Logit

$$\sim L\left(0, \frac{\pi^2}{3}\right)$$





- El marco de un umbral es provisto dentro del Modelo Probit
- La metodología incluye una variable latente (no observable) para una variable dependiente  $y^*$ , el cual puede tomar dos valores a través de una variable latente ( $y_i = 1/0$ )

- La variable dependiente del modelo es expresada como:

$$y_i^* = X_i' \beta + \mu_i$$

- Donde:

$$\mu_i \sim N(0, \sigma^2)$$

$$y_i^* \sim N(X_i' \beta, \sigma^2)$$

- ▶ Si  $y_i^* > 0$  entonces:  $y_i = 1$
- ▶ Si  $y_i^* \leq 0$  entonces:  $y_i = 0$

- Nota:

$$\text{prob}[y_i = 1] = \text{prob}[y_i^* > 0]$$



- Usando y considerando :  $prob[y_i^* > 0]$
- Restando el promedio de  $y_i^*$  de ambos lados:

$$prob[y_i^* > 0] = prob[y_i^* - X_i^* \beta > -X_i^* \beta]$$

- Dividir ambos lados respecto a los terminos de sigma  $\sigma$

$$\begin{aligned} &= prob\left[\frac{y_i^* - X_i' \beta}{\sigma} > -\frac{X_i' \beta}{\sigma}\right] \\ &= prob\left[\frac{\mu_i}{\sigma} > -\frac{X_i' \beta}{\sigma}\right] \end{aligned}$$

- Dado la **simetría** natural de la distribución, se puede reescribir la expresión:

$$= prob \left[ \frac{\mu_i}{\sigma} \leq \frac{X_i' \beta}{\sigma} \right]$$

- Nosotros reexpresamos usando el operador CDF dado por  $F(\circ)$  como:

$$= F \left[ \frac{X_i' \beta}{\sigma} \right]$$

$$= \Phi \left( \frac{X_i' \beta}{\sigma} \right)$$

**Nota** Bajo el supuesto de identificación de los parámetros del modelo probit  $\sigma = 1$

$$= \Phi \left[ X_i' \beta \right]$$

- Donde  $\Phi(\bullet)$  denotado la normal estandar de la CDF.

Areas under the  
standard normal curve



z	Second decimal place in z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000 <sup>†</sup>									

<sup>†</sup> For  $z \geq 3.9$ , the areas are 1.0000 to four decimal places.

Areas under the standard normal curve



z	Second decimal place in z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000 <sup>†</sup>									

<sup>†</sup> For  $z \geq 3.9$ , the areas are 1.0000 to four decimal places.

- Asumiendo una variable aleatoria

$$X \sim N(\mu, \sigma^2)$$

- Entonces:

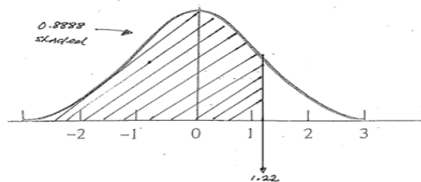
$$\frac{X_i - \mu}{\sigma} = z_i$$

- Donde:

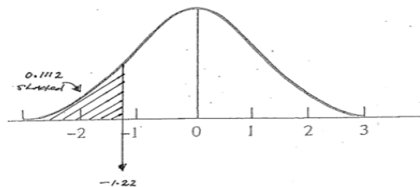
$$z_i \sim N(0, 1)$$

- ¿Cuál es el área bajo la curva normal estándar desde  $-\infty$  hasta el valor  $z$  de 1.22?

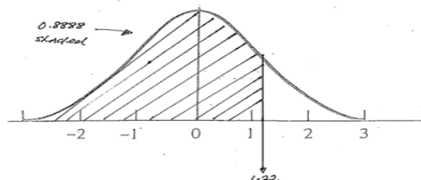
The case for  $z = 1.22$



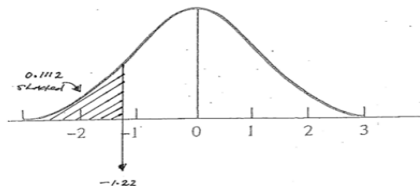
The case for  $z = -1.22$



The case for  $z = 1.22$



The case for  $z = -1.22$

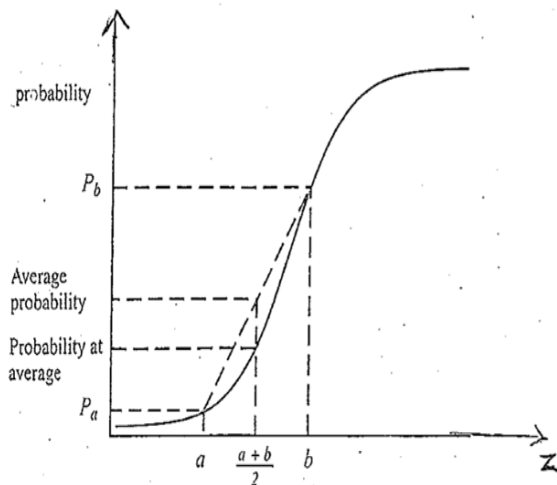


- Nosotros usamos  $\Phi(\bullet)$  denota la función de distribución acumulada para la normal estandar
- Si  $z = 1.22$ , entonces:  
 $\Phi(1.22) = 0.8888$
- Si  $z = -1.22$ , entonces:  $\Phi(-1.22)$   
 $= 1 - \Phi(1.22)$

$$= 1 - 0.8888 \Rightarrow 0.1112$$

- El area debajo la curva normal estandar de  $-\infty$  hacia 1.22 es 0.8888
- El area debajo la curva normal estandar de  $-\infty$  hacia  $-1.22$  es 0.1112

- Dado la naturaleza no lineal de la curva CDF implementada en el modelo Probit, **no se pueden usar** los valores promedio de las variables explicativas para calcular el valor promedio de **la probabilidad del producto** (como si se hace en el **modelo MCO**).
- Predecir los coeficientes del probit para los valores promedios de las variables explicativas, no retorna los valores promedio de probabilidad.





Modelo	Interpretación de coeficientes
Modelo de Probabilidad Lineal	El efecto de un pequeño cambio en el regresor sobre <b>la probabilidad asociado al evento de ocurrencia</b>
Modelo logit	El efecto de un pequeño cambio en el regresor sobre <b>el log odds ratio asociado al evento de ocurrencia</b>
Modelo Probit	El efecto de un pequeño cambio en el regresor sobre <b>índice estandarizado probit asociado al evento de ocurrencia</b>  la unidad de medida para un modelo estimado probit es una <b>desviación estandar</b>

- **Variable continua:**

$$Prob[y_i = 1] = \rho_i = \Phi(X_i' \beta)$$

- Donde, podemos definir :  $z = X' \beta$  entonces:  $\Phi(z_i)$

- Usando la regla de la cadena:

$$\frac{\partial \rho}{\partial X_k} = \frac{\partial \rho}{\partial z} \frac{\partial z}{\partial X_k}$$

- Dado:

$$\frac{\partial \rho}{\partial z} = \frac{\partial \Phi}{\partial z} = \phi(z)$$

- Luego:

$$\frac{\partial z}{\partial X_k} = \beta_k$$

- Representandolo a través de la teoria

$$\frac{\partial Prob(y_i = 1)}{\partial X_k} = \frac{\partial F}{\partial X_k} = f(X_i' \beta) \cdot \beta_k = \phi(z) \cdot \beta_k$$

- **Variable discreta:**

$$Prob[y_i = 1] = \rho_i = \Phi(X_i' \beta + \gamma D_i)$$

- Donde  $\Phi(\bullet)$ , es la CDF operador para la normal estandar
- Dado el  $D_i = 1$  si el individuo es hombre,  $D_i = 0$  es mujer
- Si,  $D_i = 1$ , la probabilidad del evento de ocurrencia es:

$$Prob[y_i = 1 | X_i, D_i = 1] = \Phi(X_i' \beta + \gamma) = \Delta_1$$

- Si,  $D_i = 0$ , la probabilidad del evento de ocurrencia es:

$$Prob[y_i = 1 | X_i, D_i = 0] = \Phi(X_i' \beta) = \Delta_0$$

- El efecto impacto del genero sobre el producto de interes es:

$$\Delta = \Phi(X_i' \beta + \gamma) - \Phi(X_i' \beta)$$

$$\Delta = \Delta_1 - \Delta_0$$

- Efecto de impacto:

$$z_i = \alpha + \beta X_i + \gamma D_i$$

- Dado  $D_i$  es una variable dummy (1/0)

