

Profesor Coordinador: Mario Castillo

Profesores: Astrid Bernal, Carlos Castellanos, Fabio Lagos, María Alejandra López, Gonzalo Torres, Hernando Mutis.

Segundo semestre de 2015

**Complementaria 11**  
**Propiedades de los estimadores.**  
**Distribuciones muestrales y uso de tablas.**

**Punto 1**

Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria de una población con distribución normal. Se cuenta con 3 estimadores para la media poblacional:

$$\hat{\mu}_1 = \frac{2(X_1 + 2X_2 + \dots + nX_n)}{n(n+1)}$$

$$\hat{\mu}_2 = \frac{X_1 + 8X_3 + 4X_7 + 5X_{10}}{18}$$

$$\hat{\mu}_3 = \frac{X_1 + X_5 + X_8 + X_{10}}{10}$$

Tenga en cuenta que el valor esperado y la varianza para una variable aleatoria con distribución normal son:

$$E(X) = \mu$$
$$VAR(X) = \sigma^2$$

- a. ¿Son insesgados los estimadores?

Para que el estimador sea insesgado, en este caso se debe cumplir que  $E(\hat{\mu}_i) = \mu$

$$E(\hat{\mu}_1) = E\left(\frac{2(X_1 + 2X_2 + \dots + nX_n)}{n(n+1)}\right)$$

$$E\left(\frac{2(X_1 + 2X_2 + \dots + nX_n)}{n(n+1)}\right) = \frac{2}{n(n+1)} E(X_1 + 2X_2 + \dots + nX_n) =$$

$$\frac{2}{n(n+1)} E\left(\sum_{i=1}^n iX_i\right) = \frac{2}{n(n+1)} \sum_{i=1}^n i E(X_i) = \frac{2}{n(n+1)} \sum_{i=1}^n i \mu$$

La fórmula de la sumatoria de n números consecutivos es:

$$\sum_{i=1}^n i = \frac{n(n+1)}{2}$$

$$E(\hat{\mu}_1) = \frac{2}{n(n+1)} * \frac{n(n+1)}{2} * \mu = \mu \rightarrow \text{insesgado}$$

Para el segundo estimador:

$$\begin{aligned} E(\hat{\mu}_2) &= E\left(\frac{X_1 + 8X_3 + 4X_7 + 5X_{10}}{18}\right) = \frac{\mu + 8\mu + 4\mu + 5\mu}{18} \\ &= \frac{18\mu}{18} = \mu \rightarrow \text{insesgado} \end{aligned}$$

Para el tercer estimador:

$$\begin{aligned} E(\hat{\mu}_3) &= E\left(\frac{X_1 + X_5 + X_8 + X_{10}}{10}\right) = \frac{\mu + \mu + \mu + \mu}{10} \\ &= \frac{4\mu}{10} = \frac{2\mu}{5} \rightarrow \text{sesgado} \end{aligned}$$

**b.** Calcule la varianza de los estimadores.

$$\begin{aligned} VAR(\hat{\mu}_1) &= VAR\left(\frac{2(X_1 + 2X_2 + \dots + nX_n)}{n(n+1)}\right) = \frac{4}{(n(n+1))^2} VAR\left(\sum_{i=1}^n ix_i\right) \\ &= \frac{4}{(n(n+1))^2} \sum_{i=1}^n i^2 VAR(x_i) = \frac{4}{(n(n+1))^2} \sigma^2 \sum_{i=1}^n i^2 \end{aligned}$$

La fórmula de la sumatoria de n números consecutivos elevados al cuadrado es:

$$\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

$$VAR(\hat{\mu}_1) = \frac{4}{(n(n+1))^2} * \sigma^2 * \frac{n(n+1)(2n+1)}{6} = \frac{2(2n+1)}{3n(n+1)} \sigma^2$$

$$VAR(\hat{\mu}_2) = VAR\left(\frac{X_1 + 8X_3 + 4X_7 + 5X_{10}}{18}\right) = \frac{106}{324} \sigma^2 = \frac{53}{62} \sigma^2$$

$$VAR(\widehat{\mu}_3) = VAR\left(\frac{X_1 + X_5 + X_8 + X_{10}}{10}\right) = \frac{4}{100} \sigma^2 = \frac{1}{25} \sigma^2$$

- c. Calcule el error cuadrático medio de los estimadores.

$$ECM(\widehat{\mu}_i) = VAR(\widehat{\mu}_i) + sesgo^2$$

Estimador 1:

$$sesgo1 = \mu - \mu = 0$$

$$ECM(\widehat{\mu}_1) = \frac{2(2n+1)}{3(n+1)} \sigma^2$$

Estimador 2:

$$sesgo2 = \mu - \mu = 0$$

$$ECM(\widehat{\mu}_2) = \frac{53\sigma^2}{62}$$

Estimador 3:

$$sesgo3 = \mu - \frac{2\mu}{5} = \frac{3\mu}{5}$$

$$ECM(\widehat{\mu}_3) = \frac{\sigma^2}{25} + \frac{9\mu^2}{25}$$

- d. ¿Cuál estimador escogería? Justifique su respuesta.

Se debe elegir un estimador insesgado y que tenga la menor varianza. Debido a que el estimador  $\widehat{\mu}_3$  es sesgado, se debe solo escoger entre  $\widehat{\mu}_1$  y  $\widehat{\mu}_2$ .

$$\frac{ECM(\widehat{\mu}_1)}{ECM(\widehat{\mu}_2)} < 1$$

$$\frac{\frac{2(2n+1)}{3(n+1)} \sigma^2}{\frac{53}{62} \sigma^2} < 1$$

$$\frac{124(2n+1)}{159(n+1)} < 1$$

$$124(2n+1) < 159(n+1) = 248n + 124 < 159n + 159$$

$$89n < 35 = n < \frac{35}{89}$$

Para un valor de n mayor a 1 el mejor estimador va a ser el segundo estimador  $\widehat{\mu}_2$ .

## Punto 2

En un proyecto espacial se ha descubierto que la variación de temperatura cada hora en una determinada zona del planeta Krypton se distribuye como una variable aleatoria  $X$  con la siguiente función de densidad de probabilidad:

$$f_X(x, \omega) = \begin{cases} \frac{1}{\omega} e^{-\frac{1}{\omega}x} & x \geq 0 \\ 0 & d.l.c \end{cases}$$

Para continuar con el proyecto, es necesario realizar una estimación apropiada del parámetro  $\omega$ , si se toma una muestra de  $n$  variaciones de temperatura independientes. Para esto, desarrolle los siguientes literales:

- a. Encuentre el estimador de máxima verosimilitud del parámetro  $\omega$ .

Paso 1: Plantear la función de máxima verosimilitud  $L(x, \omega)$ .

$$L(x, \omega) = \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \frac{1}{\omega} e^{-\frac{1}{\omega}x_i} = \left(\frac{1}{\omega}\right)^n e^{-\frac{1}{\omega}\sum_{i=1}^n x_i}$$

Paso 2: Sacar el logaritmo natural de la F.M.V  $\ln L(x, \omega)$ .

$$\begin{aligned} \ln L(x, \omega) &= \ln \left( \left(\frac{1}{\omega}\right)^n e^{-\frac{1}{\omega}\sum_{i=1}^n x_i} \right) = \ln \left(\frac{1}{\omega}\right)^n + \ln e^{-\frac{1}{\omega}\sum_{i=1}^n x_i} \\ &= -n \ln \omega - \frac{1}{\omega} \sum_{i=1}^n x_i \end{aligned}$$

Paso 3: Derivar el resultado del Paso 2 e igualar a 0 para despejar el parámetro.

$$\begin{aligned} \frac{\partial \ln L(x, \omega)}{\partial \omega} &= -\frac{n}{\omega} + \frac{1}{\omega^2} \sum_{i=1}^n x_i = 0 \\ \hat{\omega} &= \frac{\sum_{i=1}^n x_i}{n} = \bar{x} \end{aligned}$$

- b. Una vez definido el estimador para el parámetro  $\omega$ , el equipo del proyecto espacial tiene dudas sobre la consistencia del mismo. Pruebe si el estimador encontrado en el literal anterior es consistente.

Para que un estimador sea consistente se debe cumplir:

$$\begin{aligned} \lim_{n \rightarrow \infty} E(\hat{\omega}) &= \omega \\ \lim_{n \rightarrow \infty} VAR(\hat{\omega}) &= 0 \end{aligned}$$

$$E[\hat{\omega}] = E\left[\frac{\sum_{i=1}^n x_i}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^n x_i\right] = \frac{1}{n} \sum_{i=1}^n E[x_i] = \frac{1}{n} \sum_{i=1}^n \omega = \frac{1}{n} (n\omega) = \omega$$

$$\begin{aligned} Var[\hat{\omega}] &= Var\left[\frac{\sum_{i=1}^n x_i}{n}\right] = \frac{1}{n^2} Var\left[\sum_{i=1}^n x_i\right] = \frac{1}{n^2} \sum_{i=1}^n Var[x_i] = \frac{1}{n^2} \sum_{i=1}^n \omega^2 = \frac{1}{n^2} (n\omega^2) \\ &= \frac{\omega^2}{n} \end{aligned}$$

$$\lim_{n \rightarrow \infty} E[\hat{\omega}] = \lim_{n \rightarrow \infty} \omega = \omega$$

$$\lim_{n \rightarrow \infty} Var[\hat{\omega}] = \lim_{n \rightarrow \infty} \frac{\omega^2}{n} = 0$$

Dado que las dos condiciones se cumplen, se puede concluir que el estimador  $\hat{\omega}$  es consistente.

- c. Para las siguientes etapas del proyecto espacial es fundamental que el estimador del parámetro  $\omega$  sea eficiente. Compruebe que el estimador con el que ha venido trabajando es eficiente.

Para que el estimador sea eficiente, la varianza debe ser igual a la cota de Rao-Cramer:

$$VAR(\hat{\omega}) = C.R.C$$

$$C.R.C = \frac{1}{-nE\left[\frac{d^2}{d\omega^2} \ln(f(x, \omega))\right]}$$

$$\ln(f(x, \omega)) = \ln\left(\frac{1}{\omega} e^{-\frac{1}{\omega}x}\right) = -\ln(\omega) + \ln\left(e^{-\frac{1}{\omega}x}\right) = -\ln(\omega) - \frac{1}{\omega}x$$

$$\frac{\partial \ln f(x, \omega)}{\partial \omega} = -\frac{1}{\omega} + \frac{x}{\omega^2}$$

$$\frac{\partial^2 \ln f(x, \omega)}{\partial \omega^2} = \frac{1}{\omega^2} - \frac{2x}{\omega^3}$$

$$E\left(\frac{1}{\omega^2} - \frac{2x}{\omega^3}\right) = \frac{1}{\omega^2} - \frac{2E(x)}{\omega^3} = \frac{1}{\omega^2} - \frac{2\omega}{\omega^3} = -\frac{1}{\omega^2}$$

$$C.R.C = \frac{1}{-n\left(-\frac{1}{\omega^2}\right)} = \frac{\omega^2}{n}$$

Dado que la varianza del estimador  $\hat{\omega}$  es igual a la C.R.C, el estimador  $\hat{\omega}$  es eficiente, es decir, tiene la mínima varianza posible.

### Punto 3

- a. Determine la distribución de probabilidad de los siguientes estadísticos:

$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$  para una población  $X$  con distribución normal de media  $\mu$  y varianza  $\sigma^2$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$\frac{\bar{X} - \mu}{s/\sqrt{n}}$  para una población  $X$  con distribución normal de media  $\mu$  y varianza  $\sigma^2$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \quad (n-1) \frac{S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$\frac{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1) \frac{S^2}{\sigma^2}}{(n-1)}}} = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{(n-1)}$$

$(n-1) \frac{S_X^2}{\sigma^2} + (m-1) \frac{S_Y^2}{\sigma^2}$  para dos poblaciones  $X, Y$  con distribución normal de varianza  $\sigma^2$

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2$$

$$S_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{Y})^2$$

$$\frac{x_i - \mu_X}{\sigma_X} \sim N(0,1) \quad \frac{y_i - \mu_Y}{\sigma_Y} \sim N(0,1)$$

$$\left(\frac{x_i - \mu_X}{\sigma_X}\right)^2 \sim \chi^2_{(1)} \quad \left(\frac{y_i - \mu_Y}{\sigma_Y}\right)^2 \sim \chi^2_{(1)}$$

$$\sum \left(\frac{x_i - \mu_X}{\sigma_X}\right)^2 \sim \chi^2_{(n)} \quad \sum \left(\frac{y_i - \mu_Y}{\sigma_Y}\right)^2 \sim \chi^2_{(m)}$$

$$\sum \left(\frac{x_i - \bar{X}}{\sigma_X}\right)^2 \sim \chi^2_{(n-1)} \quad \sum \left(\frac{y_i - \bar{Y}}{\sigma_Y}\right)^2 \sim \chi^2_{(m-1)}$$

$$(n-1) \frac{S_X^2}{\sigma_X^2} = \frac{\sum (x_i - \bar{X})^2}{\sigma_X^2} \sim \chi^2_{(n-1)} \quad (m-1) \frac{S_Y^2}{\sigma_Y^2} = \frac{\sum (y_i - \bar{Y})^2}{\sigma_Y^2} \sim \chi^2_{(m-1)}$$

$$\chi^2_{(n-1)} + \chi^2_{(m-1)} \sim \chi^2_{(m-1)+(n-1)}$$

$\frac{S_X^2 \sigma_Y^2}{S_Y^2 \sigma_X^2}$  para dos poblaciones X,Y con distribución normal de varianza  $\sigma_X^2$  y  $\sigma_Y^2$

$$(n-1) \frac{S_X^2}{\sigma_X^2} \sim \chi^2_{(n-1)} \quad (m-1) \frac{S_Y^2}{\sigma_Y^2} \sim \chi^2_{(m-1)}$$

$$\frac{(m-1) \frac{S_Y^2}{\sigma_Y^2} / (m-1)}{(n-1) \frac{S_X^2}{\sigma_X^2} / (n-1)} = \frac{S_Y^2 \sigma_X^2}{S_X^2 \sigma_Y^2} \sim F_{(m-1, n-1)}$$

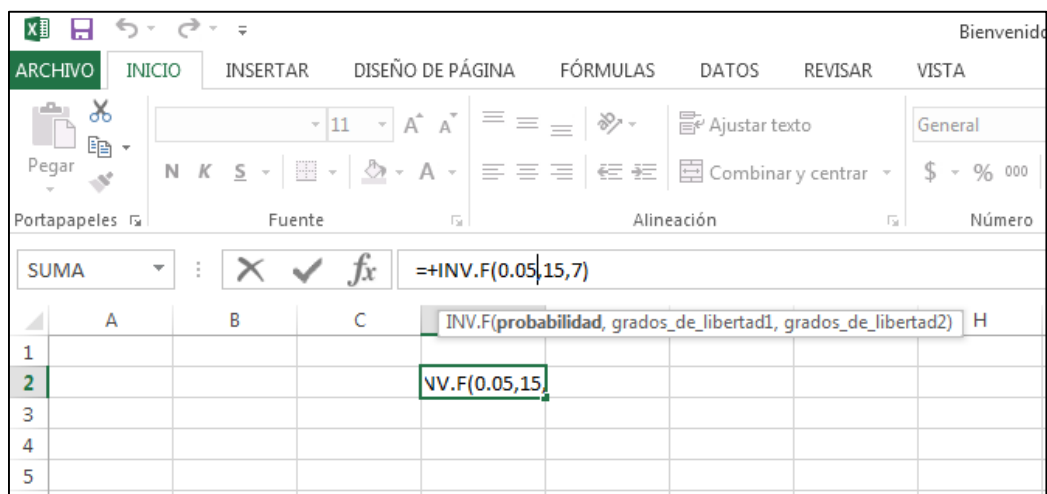
**b.** Complete la siguiente tabla con base en la información suministrada:

Distribución	Parámetro(s)	Valor de la VA	Probabilidad Acumulada hasta el valor de VA	Notación
T-Student	GL=25	a=1.7081	0.95	$P(t_{25} \leq a) = 0.95$
T-Student	GL=12	b=2.1788	0.975	$P(t_{12} \leq b) = 0.975$

F-Snedecor	15 GL en el numerador, 7 GL en el denominador	c=3.511	0.95	$P(F_{15,7} \leq c) = 0.95$
F-Snedecor	15 GL en el numerador, 7 GL en el denominador	d=0.369	0.05	$P(F_{15,7} \leq d) = 0.05$
Chi-cuadrado	Grados de libertad=19	e=21.689	0.7	$P(\chi^2_{(19)} \leq e) = 0.7$
Chi-cuadrado	GL=70	f=95.02	0.975	$P(\chi^2_{(70)} \leq f) = 0.975$
Chi-cuadrado	GL=70	g=48.75	0.025	$P(\chi^2_{(70)} \leq g) = 0.025$
T-Student	GL=20	h=-1.325	0.1	$P(t_{20} \leq h) = 0.1$

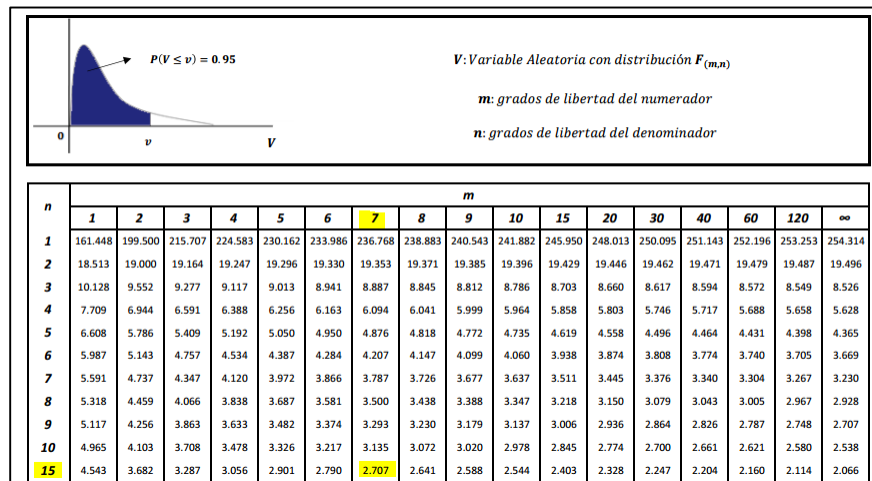
A continuación se muestra a modo de ejemplo, cómo calcular en Excel los valores de las V.A.  $d$ ,  $g$  y  $b$  de la tabla anterior, que corresponden a las distribuciones F-Snedecor, Chi-cuadrado y T-Student, respectivamente.

Para la distribución F-Snedecor, con 15 grados de libertad en el numerador y 7 en el denominador, la probabilidad acumulada hasta 0.05 se calcula mediante la función:





Mediante el uso de las tablas :



$$\frac{1}{F_{(7,15)}} = \frac{1}{2.707} = 0.369$$

Para la distribución Chi-cuadrado, con 70 grados de libertad y probabilidad acumulada hasta 0.025 se calcula mediante la función:

The screenshot displays the Microsoft Excel interface. The ribbon at the top includes tabs for ARCHIVO, INICIO, INSERTAR, DISEÑO DE PÁGINA, FORMULAS, DATOS, REVISAR, and VISTA. The 'FORMULAS' tab is active. The formula bar shows the formula `=+INV.CHICUAD(0.025,70)`. Below the formula bar, the spreadsheet grid is visible. Cell B2 is selected and contains the formula `=+INV.CHICUAD(0.025,70)`. A tooltip for the `INV.CHICUAD` function is shown, indicating its syntax: `INV.CHICUAD(probabilidad, grados_de_libertad)`. The grid also shows cells A1, B1, C1, D1, E1, F1, G1, and H1, which are currently empty.

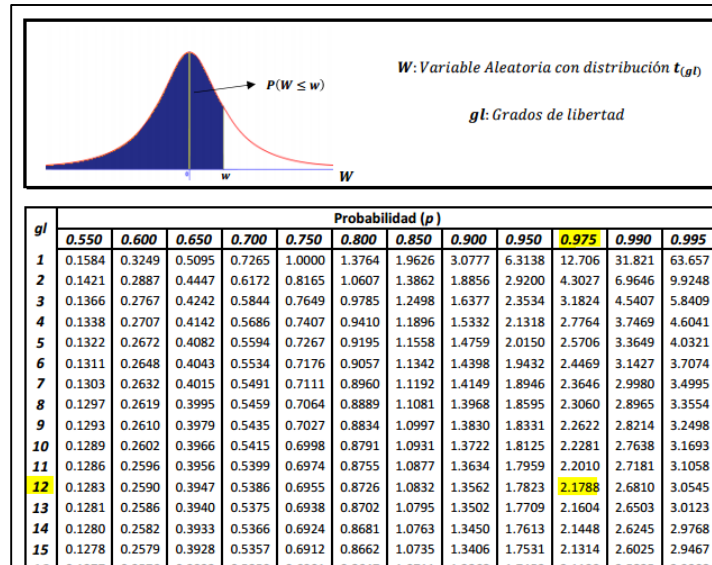
Mediante el uso de las tablas :

gl	Probabilidad (p)								
	0.005	0.010	0.025	0.050	0.100	0.200	0.250	0.300	0.400
1	0.0000	0.0002	0.0010	0.0039	0.0158	0.0642	0.1015	0.1485	0.2750
2	0.0100	0.0201	0.0506	0.1026	0.2107	0.4463	0.5754	0.7133	1.0217
3	0.0717	0.1148	0.2158	0.3518	0.5844	1.0052	1.2125	1.4237	1.8692
4	0.2070	0.2971	0.4844	0.7107	1.0636	1.6488	1.9226	2.1947	2.7528
5	0.4117	0.5543	0.8312	1.1455	1.6103	2.3425	2.6746	2.9999	3.6555
6	0.6757	0.8721	1.2373	1.6354	2.2041	3.0701	3.4546	3.8276	4.5702
7	0.9893	1.2390	1.6899	2.1673	2.8331	3.8223	4.2549	4.6713	5.4932
8	1.3444	1.6465	2.1797	2.7326	3.4895	4.5936	5.0706	5.5274	6.4226
9	1.7349	2.0879	2.7004	3.3251	4.1682	5.3801	5.8988	6.3933	7.3570
10	2.1559	2.5582	3.2470	3.9403	4.8652	6.1791	6.7372	7.2672	8.2955
11	2.6032	3.0535	3.8157	4.5748	5.5778	6.9887	7.5841	8.1479	9.2373
12	3.0738	3.5706	4.4038	5.2260	6.3038	7.8073	8.4384	9.0343	10.1820
13	3.5650	4.1069	5.0088	5.8919	7.0415	8.6339	9.2991	9.9257	11.1291
14	4.0747	4.6604	5.6287	6.5706	7.7895	9.4673	10.1653	10.8215	12.0785
15	4.6009	5.2293	6.2621	7.2609	8.5468	10.3070	11.0365	11.7212	13.0297
16	5.1422	5.8122	6.9077	7.9616	9.3122	11.1521	11.9122	12.6243	13.9827
17	5.6972	6.4078	7.5642	8.6718	10.0852	12.0023	12.7919	13.5307	14.9373
18	6.2648	7.0149	8.2307	9.3905	10.8649	12.8570	13.6753	14.4399	15.8932
19	6.8440	7.6327	8.9065	10.1170	11.6509	13.7158	14.5620	15.3517	16.8504
20	7.4338	8.2604	9.5908	10.8508	12.4426	14.5784	15.4518	16.2659	17.8088
21	8.0337	8.8972	10.2829	11.5913	13.2396	15.4446	16.3444	17.1823	18.7683
22	8.6427	9.5425	10.9823	12.3380	14.0415	16.3140	17.2396	18.1007	19.7288
23	9.2604	10.1957	11.6886	13.0905	14.8480	17.1865	18.1373	19.0211	20.6902
24	9.8862	10.8564	12.4012	13.8484	15.6587	18.0618	19.0373	19.9432	21.6525
25	10.5197	11.5240	13.1197	14.6114	16.4734	18.9398	19.9393	20.8670	22.6156
30	13.7867	14.9535	16.7908	18.4927	20.5992	23.3641	24.4776	25.5078	27.4416
40	20.7065	22.1643	24.4330	26.5093	29.0505	32.3450	33.6603	34.8719	37.1340
50	27.9907	29.7067	32.3574	34.7643	37.6886	41.4492	42.9421	44.3133	46.8638
60	35.5345	37.4849	40.4817	43.1880	46.4589	50.6406	52.2938	53.8091	56.6200
70	43.2752	45.4417	48.7576	51.7393	55.3289	59.8978	61.6983	63.3460	66.3961
80	51.1719	53.5401	57.1532	60.3915	64.2778	69.2069	71.1445	72.9153	76.1879

Para la distribución T-Student, con 12 grados de libertad y valor de la V.A igual a 2.1788, se calcula mediante la función:

Bienvenido						
ARCHIVO INICIO INSERTAR DISEÑO DE PÁGINA FÓRMULAS DATOS REVISAR VISTA						
<div> <div>Pegar</div> <div> <div>11</div> <div>A<sup>+</sup></div> <div>A<sup>-</sup></div> </div> <div> <div>N</div> <div>K</div> <div>S</div> </div> <div> <div>A</div> <div>A</div> </div> <div> <div>Ajustar texto</div> </div> <div> <div>Combinar y centrar</div> </div> <div> <div>\$</div> <div>%</div> <div>000</div> </div> </div>						
<div> <div>SUMA</div> <div>:</div> <div> <div>X</div> <div>✓</div> <div>fx</div> </div> <div>=+DISTR.T.N(2.1788,12,1)</div> </div>						
	A	B	C	DISTR.T.N(x, grados_de_libertad, acumulado)	G	H
1						
2				1788,12,1)		
3						
4						
5						
6						
7						

Mediante el uso de las tablas :



En resumen, en Excel las funciones empleadas son:

- =DISTR.CHICUAD(x, grados de libertad)
- =INV.CHICUAD(probabilidad, grados de libertad)
- =DISTR.F(x, grados de libertad 1, grados de libertad 2)
- =INV.F(probabilidad, grados de libertad 1, grados de libertad 2)
- =INV.T(Probabilidad,grados de libertad)
- =DISTR.T.N(x, grados de libertad, acumulado)