Requerimientos del Programa 4

Utilizando el **PSP 2.1** escribe un programa que:

- Lea del teclado dos datos:
  - *x* (número real mayor o igual a cero)
  - *dof* (número entero mayor a cero)
- Calcule *p = t(x, dof)* integrando de 0 a *x* la "*distribución t*" con *dof* grados de libertad. Para calcular la integral se utilizará el método de Simpson.
- Escriba en pantalla estos dos valores leídos y el valor calculado, de acuerdo con el siguiente formato:

```
x   = x.xxxxx
dof = xx
p   = x.xxxxx
```

NOTA:

✓ Los valores de *x* y *p* se desplegarán con 5 decimales (redondeados hacia arriba en su último dígito, por ejemplo: 0.123455 se desplegará como 0.12346, mientras que 0.123454 se desplegará como 0.12345)

Otras características que ***debe*** cumplir el programa:

- No utilizará ningún GUI para operar (funcionará desde la consola)
- Debe estar construido con programación orientada a objetos
- Debe contar con al menos 3 clases "relevantes" (la clase que contiene el "main" se cuenta como una de estas 3 clases)
- No pueden utilizarse funciones o librerías ya desarrolladas para calcular "Gamma" ni la "distribución t"
- El ***único*** código que puede ser reutilizado es el de tus programas 1 a 3
- Debe manejar apropiadamente (no tronar) ***todas*** las condiciones normales y de excepción
- Debe pasar exitosamente ***todos*** los casos de prueba (*error máximo 0.0001*):
  - Los diseñados por ti en la fase de diseño, y
  - Los siguientes 3 casos de prueba (es obligatorio incluirlos en el Diseño de las Pruebas):

Objetivo de la prueba	Instrucciones y datos de entrada	Resultados Esperados
Probar con datos correctos	Teclear en pantalla: 1.1 9	x = 1.10000 dof = 9 p = 0.35006
Probar con datos correctos	Teclear en pantalla: 1.1812 10	x = 1.18120 dof = 10 p = 0.36757
Probar con datos correctos	Teclear en pantalla: 2.75 30	x = 2.75000 dof = 30 p = 0.49500

Fin de los requerimientos

Explicación y ejemplo de cómo se realizan los cálculos (no son requerimientos)

(Tomado del curso original del PSP del Software Engineering Institute)

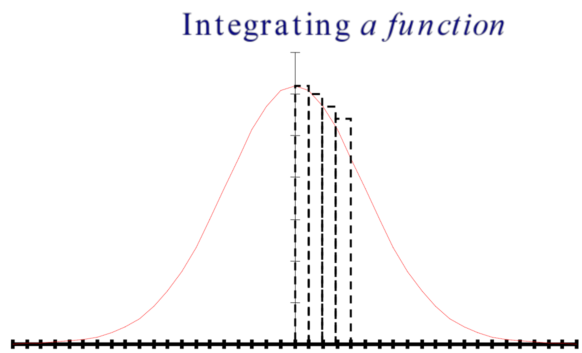
Numerical integration with Simpson's rule

Overview

Numerical integration is the process of determining the area "under" some function.

Numerical integration calculates this area by dividing it into vertical "strips" and summing their individual areas.

The key is to minimize the error in this approximation.



Simpson's rule

Simpson's rule can be used to integrate a symmetrical statistical distribution function over a specified range (e.g., from 0 to some value *x*).

1. *num_seg* = initial number of segments, an even number
2. W = *x*/*num_seg*, the segment width
3. E = the acceptable error, e.g., 0.0000001
4. Compute the integral value with the following equation.

$$p = \frac{W}{3}\left[F(0) + \sum_{i=1,3,5\ldots}^{num\_seg-1} 4F(iW) + \sum_{i=2,4,6\ldots}^{num\_seg-2} 2F(iW) + F(x)\right]$$

5. Compute the integral value again, but this time with *num_seg* = *num_seg**2.
6. If the difference between these two results is greater than E, double *num_seg* and compute the integral value again. Continue doing this until the difference between the last two results is less than E. The latest result is the answer.

Continued on next page

Numerical integration with Simpson's rule, Continued

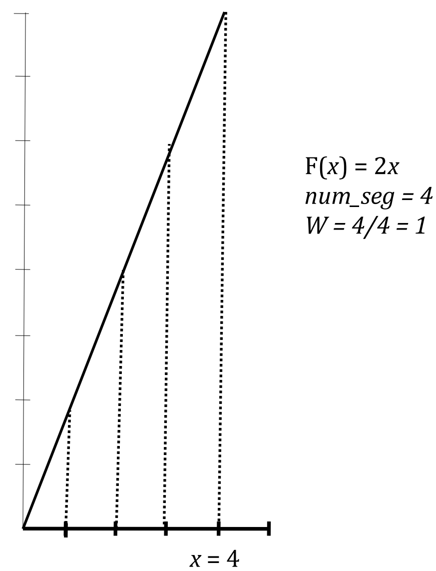
A simple example

Let's look at a simple function, where F(x) = 2x.

Note: This example is a triangle. The area of a triangle is

$$\frac{1}{2}(base)(height)$$

$$\frac{1}{2}(4)(8) = \frac{32}{2} = 16$$



In this example, we can expand Simpson's rule

$$p = \frac{W}{3}\left[F(0) + \sum_{i=1,3,5...}^{num\_seg-1} 4F(iW) + \sum_{i=2,4,6...}^{num\_seg-2} 2F(iW) + F(x)\right]$$

to

$$p = \frac{1}{3}\left[F(0) + 4F(1) + 2F(2) + 4F(3) + F(4)\right]$$

and then substitute calculated values for the function F(x)= 2x

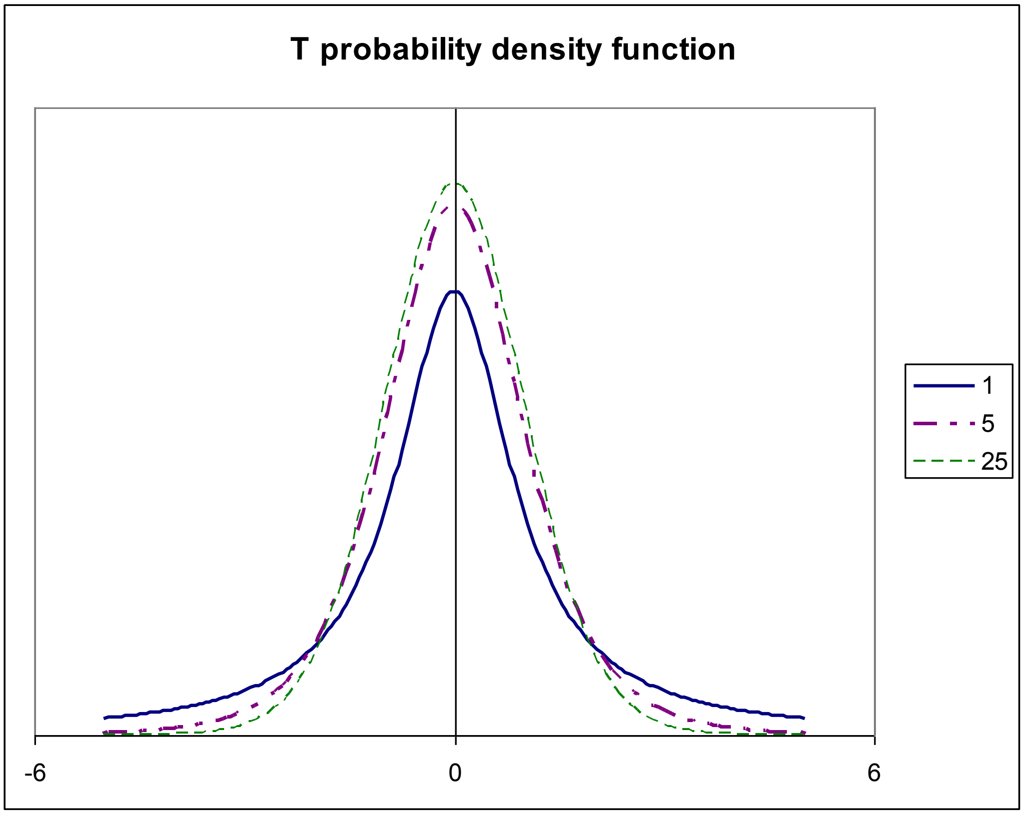
$$p = \frac{1}{3}\left[(0) + 4(2) + 2(4) + 4(6) + (8)\right] = \frac{1}{3}\left[0 + 8 + 8 + 24 + 8\right] = \frac{48}{3} = 16$$

The t distribution

Overview

The t distribution is a very important statistical tool. It is used instead of the normal distribution when the true value of the population variance is not known and must be estimated from a sample.

The shape of the t distribution is dependent on the number of points in your dataset. As *n* gets large, the t distribution approaches the normal distribution. For lower values, it has a lower central "hump" and fatter "tails."



Using the t distribution in the PSP

In the PSP the t distribution is used in two ways. We use the t distribution to test the significance of a correlation. We also use the t distribution to calculate the prediction interval when using PROBE methods A and B.

Continued on next page

T distribution function

When numerically integrating the t distribution with Simpson's rule, use the following function.

$$F(x) = \frac{\Gamma\left(\frac{dof+1}{2}\right)}{(dof * \pi)^{1/2}\,\Gamma\left(\frac{dof}{2}\right)}\left(1+\frac{x^2}{dof}\right)^{-(dof+1)/2}$$

where
- dof = degrees of freedom
- Γ is the gamma function

The gamma function is
$\Gamma(x) = (x-1)\Gamma(x-1)$, where
- $\Gamma(1) = 1$
- $\Gamma(1/2) = \sqrt{\pi}$

Continued on next page

The t distribution, Continued

An example of calculating gamma for an integer value

$\Gamma(x)$ for integer values is $\Gamma(x) = (x-1)!$.

$\Gamma(5) = 4! = 24$

An example of calculating gamma for a non-integer value

$$\Gamma\left(\frac{9}{2}\right) = \frac{7}{2}\Gamma\left(\frac{7}{2}\right)$$

$$\frac{7}{2}\Gamma\left(\frac{7}{2}\right) = \frac{7}{2}*\frac{5}{2}\Gamma\left(\frac{5}{2}\right)$$

$$\frac{7}{2}*\frac{5}{2}\Gamma\left(\frac{5}{2}\right) = \frac{7}{2}*\frac{5}{2}*\frac{3}{2}\Gamma\left(\frac{3}{2}\right)$$

$$\frac{7}{2}*\frac{5}{2}*\frac{3}{2}\Gamma\left(\frac{3}{2}\right) = \frac{7}{2}*\frac{5}{2}*\frac{3}{2}*\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

$$\frac{7}{2}*\frac{5}{2}*\frac{3}{2}*\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{7}{2}*\frac{5}{2}*\frac{3}{2}*\frac{1}{2}*\sqrt{\pi} = 11.63173$$

An example

An example

In this example, we'll calculate the values for the t distribution integral from 0 to x =1.1 with 9 degrees of freedom.

1. First, we'll set, as an example, num_seg = 10 (can be any even number)
2. W = x/num_seg = 1.1/10 = 0.11
3. E = 0.0000001
4. dof = 9
5. x = 1.1
6. Compute the integral value with the following equation.

$$p = \frac{W}{3}\left[F(0) + \sum_{i=1,3,5...}^{num\_seg-1} 4F(iW) + \sum_{i=2,4,6...}^{num\_seg-2} 2F(iW) + F(x)\right] \text{where}$$

$$F(x) = \frac{\Gamma\left(\frac{dof+1}{2}\right)}{(dof*\pi)^{1/2}\,\Gamma\left(\frac{dof}{2}\right)}\left(1+\frac{x^2}{dof}\right)^{-(dof+1)/2}$$

7. We can solve the first part of the equation:

$$\frac{\Gamma\left(\frac{dof+1}{2}\right)}{(dof*\pi)^{1/2}\,\Gamma\left(\frac{dof}{2}\right)} = \frac{24}{5.3174*11.6317} = 0.388035$$

The intermediate values for this are in the Table 2.

i	x_i	$1+\frac{x_i^2}{dof}$	$\left(1+\frac{x_i^2}{dof}\right)^{-\left(\frac{dof+1}{2}\right)}$	$\frac{\Gamma\left(\frac{dof+1}{2}\right)}{(dof*\pi)^{1/2}\Gamma\left(\frac{dof}{2}\right)}$	$F(x_i)$	Multiplier	Terms $\frac{w}{3}*Multiplier*F(x_i)$
0	0	1	1	0.388035	0.38803	1	0.0142279467
1	0.11	1.00134	0.9933	0.388035	0.38544	4	0.0565307512
2	0.22	1.00538	0.97354	0.388035	0.37777	2	0.0277029371
3	0.33	1.0121	0.94164	0.388035	0.36539	4	0.0535901651
4	0.44	1.02151	0.89905	0.388035	0.34886	2	0.0255833110
5	0.55	1.03361	0.84765	0.388035	0.32892	4	0.0482410079
6	0.66	1.0484	0.78952	0.388035	0.30636	2	0.0224665901
7	0.77	1.06588	0.72688	0.388035	0.28205	4	0.0413680625
8	0.88	1.08604	0.66185	0.388035	0.25682	2	0.0188336386
9	0.99	1.1089	0.5964	0.388035	0.23142	4	0.0339422246
10	1.1	1.13444	0.53221	0.388035	0.20652	1	0.0075722695
Result							**0.3500589043**

Table 2

Continued on next page

Example, continued

8. Compute the integral value again, but this time with *num_seg* = 20. The new result is 0.3500586369.
9. We compare the new result to the old result.
10. | 0.3500589043 - 0.3500586369 | = 0.0000002674
11. Since 0.0000002674 > 0.0000001 (*E*), compute the integral value again, but this time with *num_seg* = 40. The new result is 0.3500586202.
12. We compare the new result to the old result.
13. | 0.3500586369 - 0.3500586202 | = 0.0000000167
14. Since 0.0000000167 < 0.0000001 (*E*), we can stop and return the last value ***p* = 0.3500586202**.

In summary, we keep computing the integral value (each time doubling the number of segments), until the absolute difference of the new result minus the old one is less than *E*.