

Praktiketan SPSSrekin

Elena Agirre
Inmaculada Lecubarri
Jose Mari Eguzkitza

Praktiketan

SPSSrekin

Elena Agirre
Inmaculada Lecubarri
Jose Mari Eguzkitza

© Elena Agirre Basurko
© Inmaculada Lecubarri Alonso
© Jose Mari Eguzkitza Arrizabalaga

ISBN: 978-84-615-7798-9
Lege-gordailua: VI-163-2012

Argitalpena:
E. Agirre, I. Lecubarri, J.M. Eguzkitza

Industria Ingeniaritza Teknikoko Unibertsitate Eskola
UPV/EHU
La Casilla, 3
48012, Bilbo

Koordinatzailea:
E.Agitre (elena.agirre@ehu.es)

Azalaren diseinua:
Albaro Anta Sanz

Inprimategia:
Printvisión., S.L.
C/ Rioja, 27
01005, Vitoria-Gasteiz

HITZAURREA

*Praktiketan SPSSrekin izeneko liburu hau gida bat izateko asmoz egin da. Norentzako gida? Euskal Herriko Unibertsitateko Bilboko Industria Ingeniaritza Teknikoko Unibertsitate-Eskolan, *Ingeniaritzaren Metodo Estatistikoak* irakasgaia ikasten ari diren ikasleentzat. Liburu hau ez da Estatistikako liburua, ezta SPSSko manuala ere. SPSS programaren bidezko ordenagailu-praktikak egiteko laguntza eman nahi du liburu honek. Horrela, ikasleak bere ikasketa-prozesuan aurrera egiteko tresna bat gehiago izango du. Bere helburua praktikoa izanik, gai bakoitzaren hasieran beharrezkoak diren azalpenak emango dira era laburrean, eta ondoren SPSS programa erabiliz ebatzitako adibide batzuk aurkeztu eta garatuko dira. Bukatzeko, gai bakoitzaren amaieran ikasleak praktikatzeko ariketak aurkituko ditu. Oro har, adibide eta ariketetako enuntziatuak Industriako Ingeniaritza Gradua eta Informazio Sistemen Informatikaren Ingeniaritzako Gradua ikasten duten ikasleentzako egokiak direlakoan gaude.*

Bilbon, 2012ko martxoan

Aurkibidea

Hitzaurrea	5
1. SPSSren hastapenak	7
1.1. Sarrera.....	7
1.2. SPSS ireki, datuak sartu, irakurri eta gorde.....	7
1.3. Aldagaien transformazioa.....	14
1.4. Adibide batzuk.....	15
1.5. Praktikatzeko ariketak.....	22
2. Estatistika deskribatzailea	23
2.1. Sarrera.....	23
2.2. Adibide batzuk.....	28
2.3. Praktikatzeko ariketak.....	50
3. Probabilitate-banaketa diskretuak	53
3.1. Aldagai aleatorio diskretuaren probabilitate-funtzioa eta banaketa-funtzioaren balioak SPSS erabiliz.....	53
3.2. Adibide batzuk.....	54
3.3. Praktikatzeko ariketak.....	69
4. Probabilitate-banaketa jarraituak	71
4.1. Aldagai aleatorio jarraituaren dentsitate-funtzioa, banaketa- funtzioa eta alderantzizko banaketa-funtzioa SPSS erabiliz.....	71
4.2. Adibide batzuk.....	73
4.3. Praktikatzeko ariketak.....	85
5. Estimazioa	87
5.1. Sarrera.....	87
5.2. Adibide batzuk.....	87
5.3. Praktikatzeko ariketak.....	105

6. Hipotesi-kontrastea	107
6.1. Zenbait kontzeptu hipotesi-kontraste parametrikoan.....	107
6.2. Adibide batzuk.....	108
6.3. Praktikatzeko ariketak.....	118
Sinboloak eta laburdurak.....	121
Bibliografia.....	123

1. SPSSren HASTAPENAK

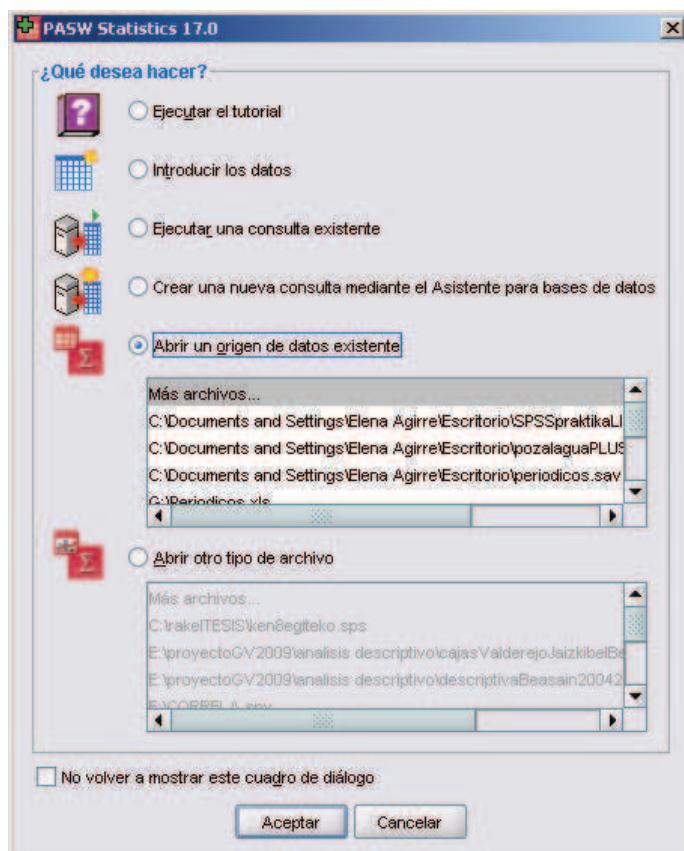
1.1. Sarrera

SPSS (*Statistical Package for Social Sciences*) programa informatikoa 1968 urtean sortu zen. Oso erabilia den software estatistikoa da, tamaina handiko datu-baseekin lan egiteko duen gaitasunagatik. SPSS programa erabiltzeko erraza da: menuan dauden prozeduren artean aukeratu botoietan klik eginez, eta kasu bakoitzean interesekoa den teknika estatistikoaren emaitzak lortuko dira taula eta grafikoen bidez. SPSS erabiliz teknika estatistiko anitz aplika daitezke: datuen kudeaketa, datuen analisi estatistikoa, probabilitateen kalkulua, estimazioa, hipotesi-kontrastea, faktore-analisia, bariantzaren analisia,

Gidaliburu honetan Industriako Ingeniaritza Gradua eta Kudeaketaren eta Informazio Sistemen Informatikaren Ingeniaritzako Gradua ikasten diharduten ikasleentzat interesgarrienak diren teknika estatistikoak aztertuko dira SPSSren PASW Statistics 17.0 bertsioa erabiliz.

1.2. SPSS ireki, datuak sartu, irakurri eta gorde

SPSS programaren 17.0 bertsioa irekitzen denean, 1.1. irudiko pantaila azaltzen da. Bertan aukera ezberdinak daude. Adibidez, SPSSko fitxategi batean dauden datuak ireki edo beste mota bateko fitxategi bat ireki daiteke.



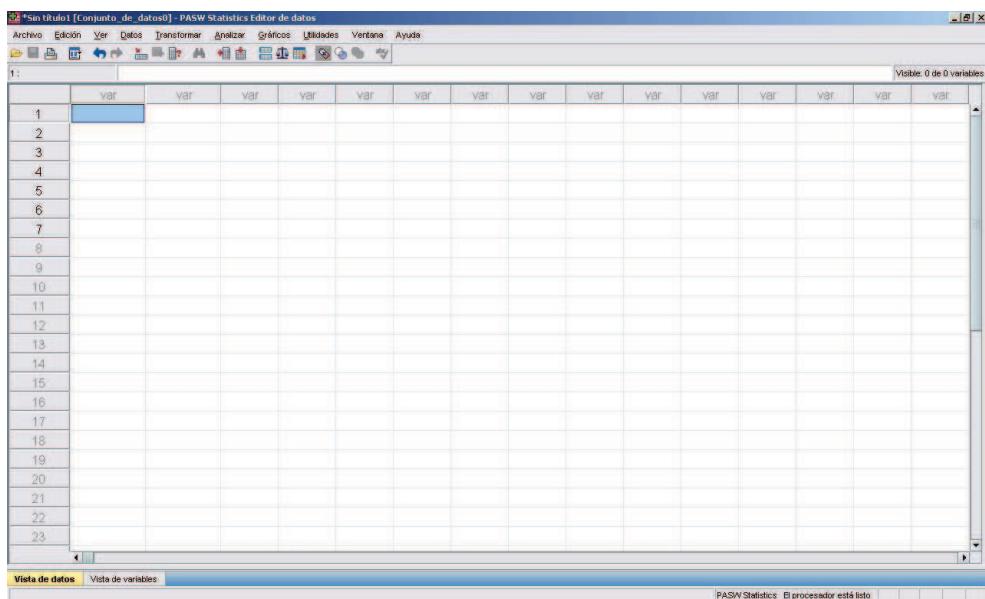
1.1. irudia.

SPSS programak hiru fitxategi-mota erabiltzen ditu:

- .sav amaierakoak, SPSSren datu-baseak dira.
- .spo amaierakoak, SPSSren grafiko eta emaitzen fitxategiak dira.
- .sps, sintaxiko fitxategiak dira.

Hasteko, datuak sartu nahi baditugu, *Introducir los datos* delakoa aukeratuko dugu.

Orduan 1.2. irudiko pantaila azalduko da. *Vista de datos* izeneko leihoko datuak sartuko dira. Aldagaiak zutabetan idazten dira eta lerroak kasuak dira.

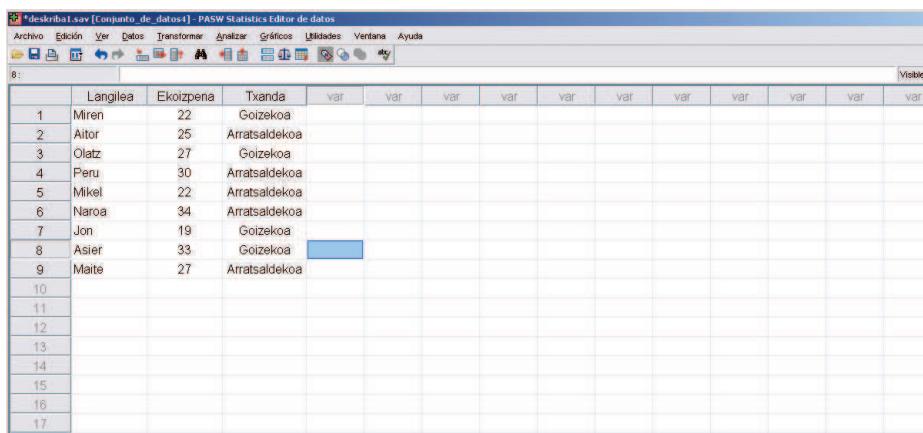


1.2. irudia.

Praktikatzeko, sar ditzagun 1.1. taulako datuak SPSSko datuen leihoa:

1.1 taula.

Langilearen izena	Eguneko ekoizpena	Txanda
Miren	22	Goizekoa
Aitor	25	Arratsaldekoa
Olatz	27	Goizekoa
Peru	30	Arratsaldekoa
Mikel	22	Arratsaldekoa
Naroa	34	Arratsaldekoa
Jon	19	Goizekoa
Asier	33	Goizekoa
Maite	27	Arratsaldekoa



	Langilea	Ekoizpena	Txanda	var										
1	Miren	22	Goizekoa											
2	Aitor	25	Arratsaldekoa											
3	Olatz	27	Goizekoa											
4	Peru	30	Arratsaldekoa											
5	Mikel	22	Arratsaldekoa											
6	Naroa	34	Arratsaldekoa											
7	Jon	19	Goizekoa											
8	Asier	33	Goizekoa											
9	Maite	27	Arratsaldekoa											
10														
11														
12														
13														
14														
15														
16														
17														

1.3. irudia.

Datuak sartuta daudelarik (1.3. irudia), aldagaien ezaugarriak ikusteko *Vista de variables* leioha (1.4. irudia) kontsultatuko da.



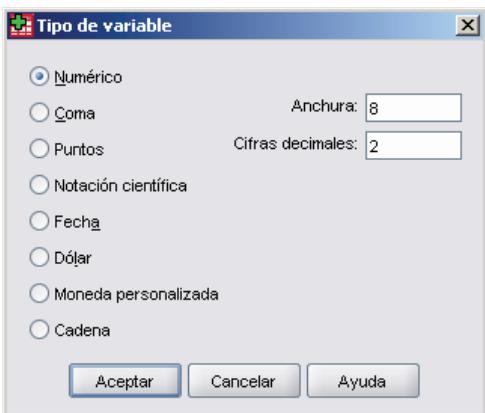
	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	Langilea	Cadena	8	0		Ninguna	Ninguna	8	Izquierda	Nominal
2	Ekoizpena	Numérico	8	0		Ninguna	Ninguna	8	Centrado	Escala
3	Txanda	Cadena	16	0		Ninguna	Ninguna	9	Centrado	Nominal
4										
5										
6										
7										
8										
9										
10										
11										
12										

1.4. irudia.

Aldagaia definitzen duten ezaugarriak hauexek dira:

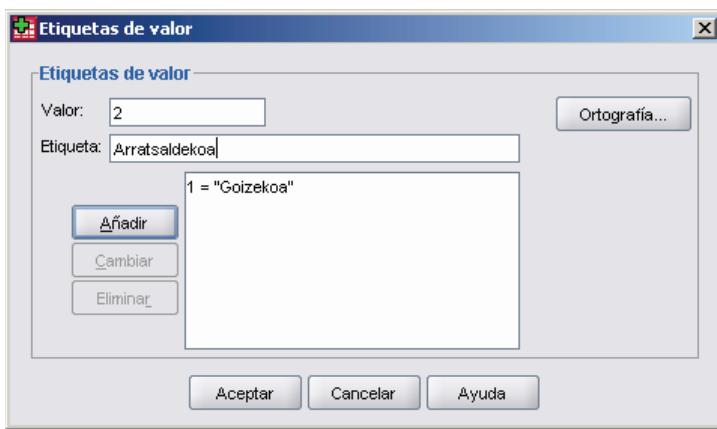
- Aldagaiaren izena:** *Nombre* izeneko zutabean adieraziko da. 1.4.irudiko adibidean hiru aldagi daude: *Langilea*, *Ekoizpena* eta *Txanda*.
- Aldagai-mota:** aldagaiaren formatua adierazteko *Tipo* izeneko zutabean dauden moten artean aukeratu behar da. Horrela, aurreko adibidean *Langilea* eta *Txanda* aldagaiak alfazebakizkoak (*cadena*) dira eta

Ekoizpena zenbakizkoa da. Hurrengo irudiak SPSS programako aldagai-motak azaltzen dira:



1.5. irudia.

3. Zabalera: *Anchura* izeneko zutabeen aldagai bakoitzari eman zaion espazioa adierazten da. Aztergai den kasuan (1.4. irudia) 8 edo 16 unitateko zabalerakoak dira aldagaiak.
4. Dezimalak: *Decimales* izenak aldagai bakoitzaren zifra dezimalen kopurua adierazten du, aldagaiaren zenbakizkoan denean. 1.4. irudiko adibidean zenbakizko aldagaiak 0 zifra dezimal ditu.
5. Etiketak: *Etiquetas* zutabeak aldagaiaren izenari buruz informazio gehiago zehaztea ahalbidetuko du.
6. Balioen etiketak: aldagai kualitatiboa edo ordinala bida, balioari etiketa jar daki. Aztergai den adibidean, *Txanda* aldagaiari 1 balioa eman zaio *Goizekoa* adierazteko eta 2 balioa ezarri da *Arratsaldekoa* denean txanda, 1.6. irudian adierazten den moduan.

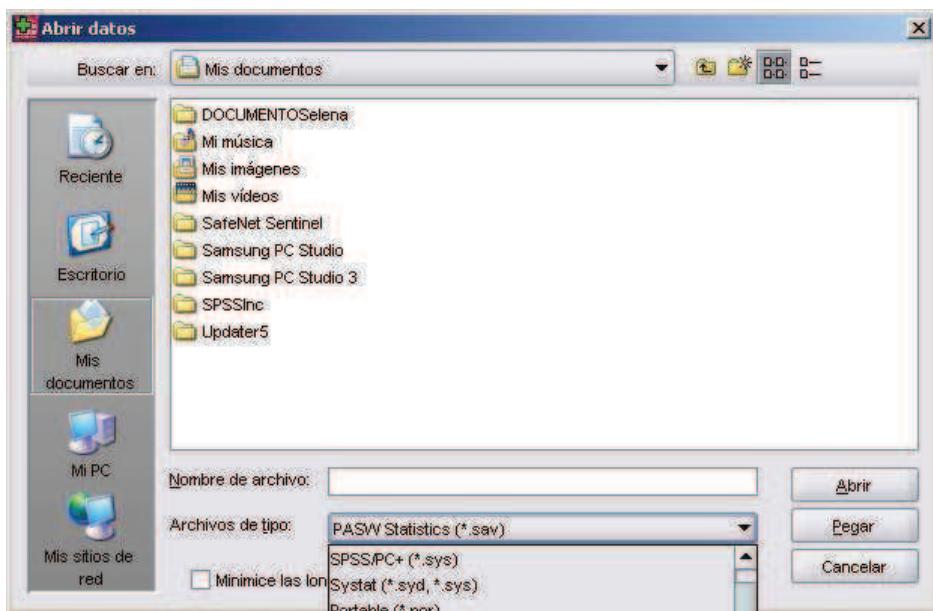


1.6. irudia.

Etiketa hauek grafikoetan eta txostenetan adieraziko dira.

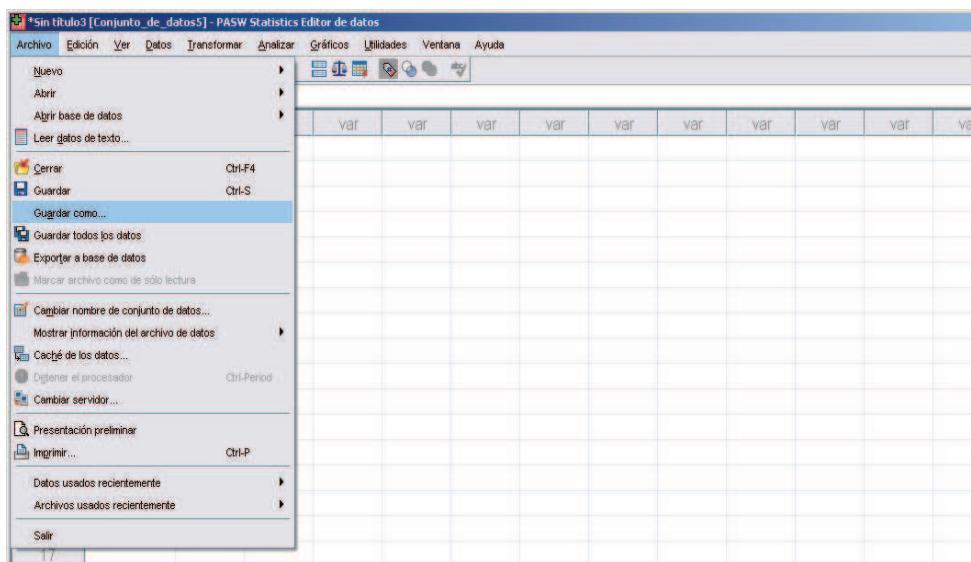
7. Balio galduak: falta diren balioak (*Perdidos*) dira. Halakoetan gelaxkak hutsik utzi daitezke ala balio zehatzekin ordezkatu.
8. Zutabeak: aldagai-motan adierazitako zabaleraren berdina edo handiagoa izan behar du *Columnas* delakoak.
9. Lerrokatzea: zutabe bakoitzeko informazioa eskuinean, ezkerrean edo erdian zentratuta dagoen adieraziko da *Alineación* zatabean.
10. Neurketa: erabilitako aldagai estatistikoa zein motatakoa den adieraziko da *Medición* zatabean. Horrela, aldagai kuantitatiboa denean *Escala* hitza adieraziko da zutabe horretan, eta aldagai kualitatiboa denean *Ordinal* edo *Nominal* hitzak izango dira.

Halaber, SPSS programaren barran *Archivo > Abrir > Datos* aukeratuz, 1.7. irudian ikus daitekeen moduan, .sav, .xls, .por, ... luzapenak dituzten mota ezberdinak fitxategiak ireki eta beraiekin lan egin daiteke.



1.7. irudia.

Bukatzeko, egindako lana gordetzeko formato ezberdinak (*Guardar como...*) aukera daitezke (1.8. irudia).



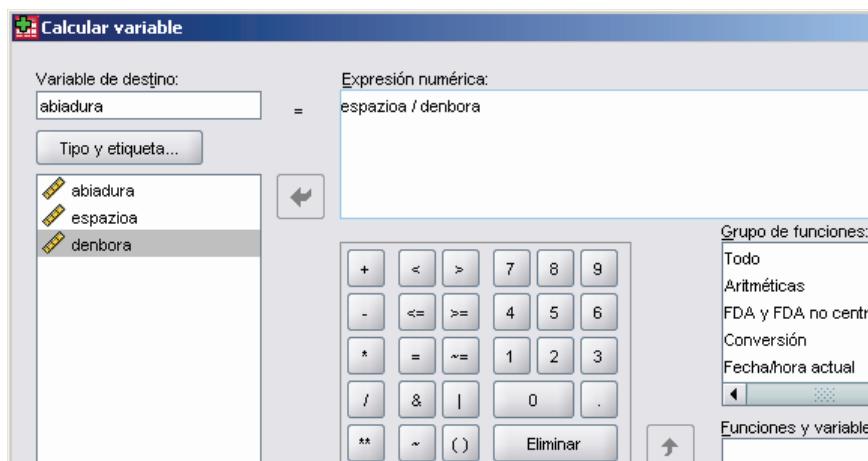
1.8. irudia.

1.3. Aldagaien transformazioa

SPSSren bidez aldagaiak transforma daitezke, aldagai berriak sortuz. Adibidez, *Transformar > Calcular variable* prozedura aukeratuz, 1.9. irudiko koadroa agertuko da, non *abiadura* izeneko aldagaia sortuko den

$$\text{abiadura} = \text{espazioa} / \text{denbora}$$

adierazpena erabiliz.



1.9. irudia.

Zenbakizko adierazpenean matematikako zenbait eragile edo funtzioidatzik dira aldagai berria lortzeko prozeduran. 1.2. taulan eragile funtsezkoenak laburbildu dira:

1.2. taula.

Eragilea	Esanahia
=	Berdin
~=	Ezberdin
<	Txikiago
>	Handiago
<=	Txikiago edo berdin
>=	Handiago edo berdin
&	Eta
	Edo
~	Ez
+	Batuketa
-	Kenketa
*	Biderketa
/	Zatiketa
**	Berreketa
<i>ABS(arg)</i>	Balio absolutua
<i>RND(arg)</i>	Biribiltzea
<i>TRUNC(arg)</i>	Zati osoa
<i>SQRT(arg)</i>	Erro karratua
<i>EXP(arg)</i>	e zenbakiaren berreketa
<i>LG10(arg)</i>	Logaritmo hamartarra
<i>LN(arg)</i>	Logaritmo nepartarra

1.4. Adibide batzuk**1.4.1. adibidea**

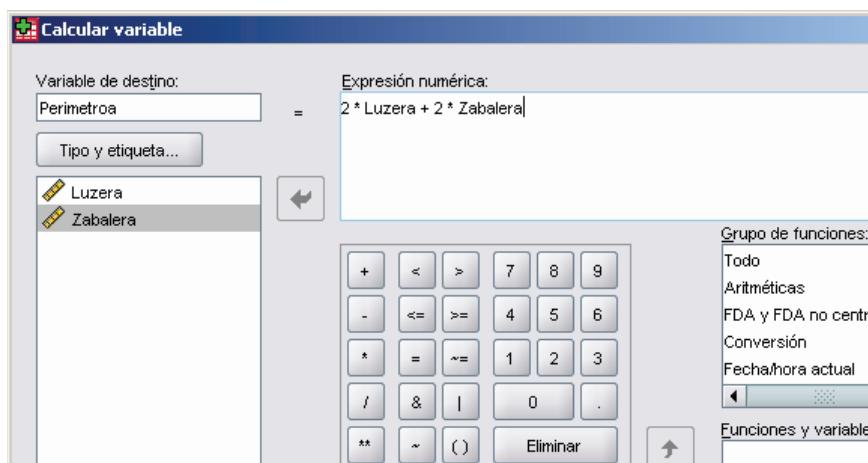
Hurrengo datuak hamar xafla errektangeluarren luzerak eta zabalerak dira zentimetrotan neurtuta:

Luzerak:	10	13	9,5	8,5	8,7	9,1	11	11,3	10,5	10
Zabalerak:	4	5	3	2,5	1,9	2,3	3	4	4,4	2,8

Sortu bi aldagai berri xafla hauen perimetroak eta azalerak neurtzeko.

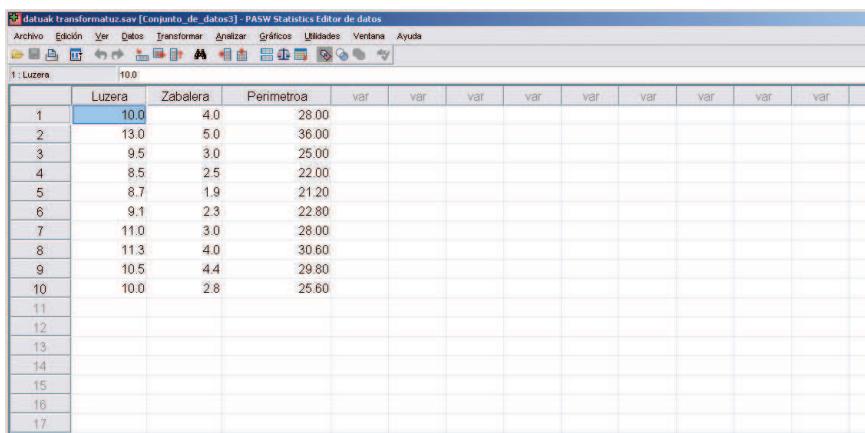
Ebazpena:

Datuak sartu eta gero, *Transformar > Calcular variable* prozedura aukeratuko da. Irekiko den leihoaan *Expresión numérica* dioen koadroan errektangeluaren perimetroaren adierazpena idatziko da (1.10 irudia).



1.10. irudia.

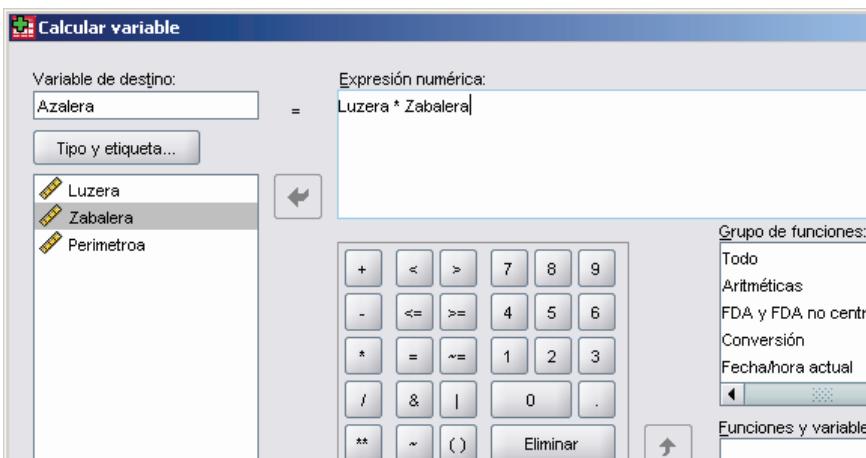
Era horretan 1.11. irudian adierazten den *Perimetroa* izeneko aldagai berria sortu da:



	Luzera	Zabalera	Perimetroa	var								
1	10.0	4.0	28.00									
2	13.0	5.0	36.00									
3	9.5	3.0	25.00									
4	8.5	2.5	22.00									
5	8.7	1.9	21.20									
6	9.1	2.3	22.80									
7	11.0	3.0	28.00									
8	11.3	4.0	30.60									
9	10.5	4.4	29.80									
10	10.0	2.8	25.60									
11												
12												
13												
14												
15												
16												
17												

1.11. irudia.

Era berean, *Transformar > Calcular variable* aukeratuz, 1.12. irudiko koadroa irekiko da:



1.12. irudia.

Agindua exekutatu ondoren, *Azalera* izeneko aldagai berria sortuko da (1.13. irudia).

1.13. irudia.

1.4.2. adibidea

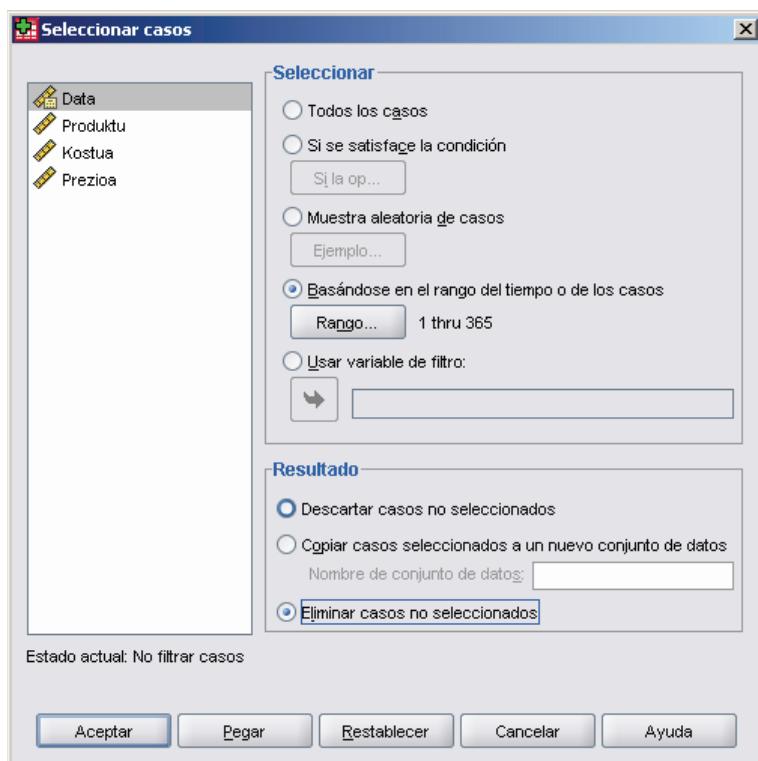
Hurrengo adibidean .sav luzapeneko datu-base bat ireki da eta bertan 2006-2007 denboraldiko egun bakoitzean enpresa batek zenbat produktu ekoitzi dituen, zein den produktuen elaborazio-kostua eta zein den salmenta-prezioa adierazi dira, *Data, Produktu, Kostua eta Prezioa* izeneko aldagaietan, hurrenez hurren (1.14. irudia).

1.14. irudia.

- Aukeratu 2006 urteko datuak eta ezabatu gainontzeko datu guztiak.
- Kalkulatu *Irabaziak* izeneko aldagai berria, prezioaren eta kostuaren arteko differentzia moduan.
- Gorde *Data* eta *Irabaziak* dituen datu-basea excel fitxategi batean.

Ebazpena:

- 2006 urteko datuak aukeratzeko emandako datu-baseko lehenengo 365 kasuak hartu behar dira. Orduan, SPSSko menuen aukeren artean *Datos > Seleccionar casos* aukeratuz gero, 1.15. irudiko koadroa irekiko da:



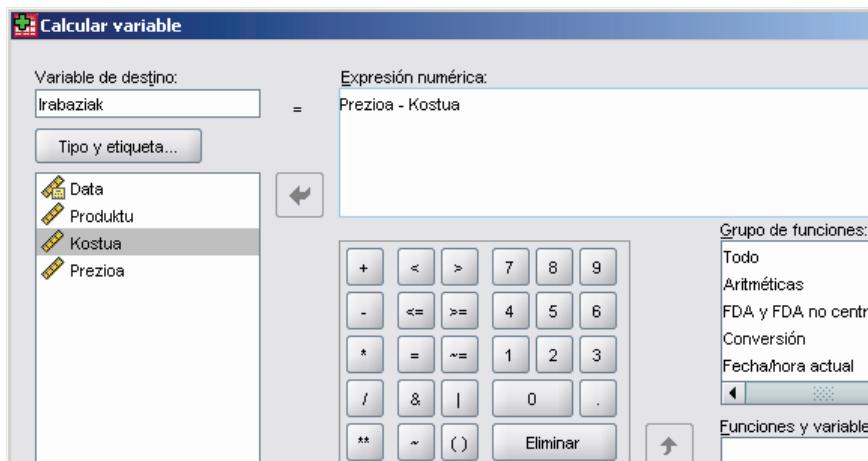
1.15. irudia.

1.15. irudian aukeratutakoa exekutatzen denean, lehenengo 365 kasuak dituen datu-basea geratuko da, aukeratu ez diren kasuak ezabatu egingo baitira.

- b) Irabaziak izeneko aldagaia lortzeko adierazpen matematikoa

Prezioa – Kostua

da. Gero, *Transformar > Calcular variable* prozedura aukeratu eta zenbakizko adierazpenean aurreko diferentzia idatziko da (1.16. irudia).



1.16. irudia.

	Data	Produktu	Kostua	Prezioa	Irabaziak	var						
1	01-Jan-2006	79	134.30	256.75	122.45							
2	02-Jan-2006	77	130.90	250.25	119.35							
3	03-Jan-2006	44	74.80	143.00	68.20							
4	04-Jan-2006	56	95.20	182.00	86.80							
5	05-Jan-2006	53	90.10	172.25	82.15							
6	06-Jan-2006	43	73.10	139.75	66.65							
7	07-Jan-2006	57	96.90	165.25	88.35							
8	08-Jan-2006	53	90.10	172.25	82.15							
9	09-Jan-2006	93	158.10	302.25	144.15							
10	10-Jan-2006	41	69.70	133.25	63.55							
11	11-Jan-2006	61	103.70	198.25	94.55							
12	12-Jan-2006	41	69.70	133.25	63.55							
13	13-Jan-2006	50	85.00	162.50	77.50							
14	14-Jan-2006	79	134.30	256.75	122.45							
15	15-Jan-2006	59	100.30	191.75	91.45							
16	16-Jan-2006	43	73.10	139.75	66.65							
17	17-Jan-2006	85	144.50	276.25	131.75							
18	18-Jan-2006	75	126.65	242.13	115.48							
19	19-Jan-2006	64	108.80	208.00	99.20							
20	20-Jan-2006	69	117.30	224.25	106.95							
21	21-Jan-2006	75	127.50	243.75	116.25							
22	22-Jan-2006	48	81.60	156.00	74.40							
23	23-Jan-2006	69	117.30	224.25	106.95							

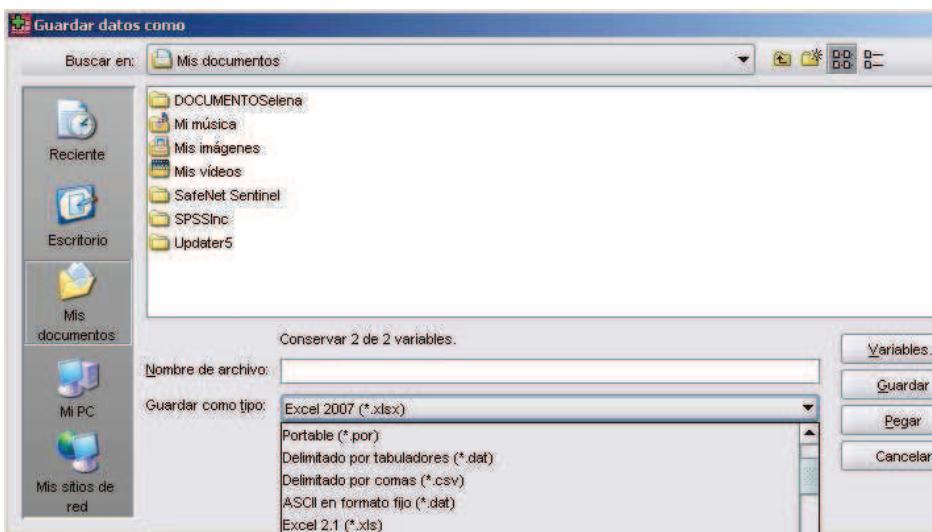
1.17. irudia.

1.17 irudian jaso dira enpresaren 2006 urteko irabaziak.

- c) Oraingoan *Data* eta *Irabaziak* aldagaiak dituen datu-basea hartuko da (1.18. irudia), eta SPSSren menuan *Archivo > Guardar como* aukeratuz, 1.19 irudiko koadroa irekiko da eta bertan Excel moduan gordetzeko aginduko dugu.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida
1	Data	Fecha	11	0		Ninguna	Ninguna	11	Derecha	Escala
2	Irabaziak	Numerico	8	2		Ninguna	Ninguna	7	Derecha	Escala
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										

1.18. irudia.



1.19. irudia.

1.5. Praktikatzeko ariketak

1.5.1. ariketa

Idatzi hurrengo taulako datuak SPSS programan eta gorde .sav fitxategi batean:

Graduatuena NA	Berezitasuna	Graduazioko batez besteko nota
16345678A	Elektrizitatea	6,5
16456781B	Elektronika	7,0
16567812C	Kimika	6,0
16312345D	Kimika	8,2
15335555A	Mekanika	5,7
16512678S	Elektronika	7,6
15556677M	Mekanika	5,5
15223344Q	Informatika	5,2
15641328N	Informatika	7,4
16354367M	Elektronika	6,3
15224569R	Elektrizitatea	6,6
15532128U	Mekanika	8,6

Sortu *Kalifikazioa* izeneko aldagai berri bat, zeinak graduazioko batez besteko notaren arabera *oso ongi, ongi* edo *nahiko* balioak jasoko dituen.

1.5.2. ariketa

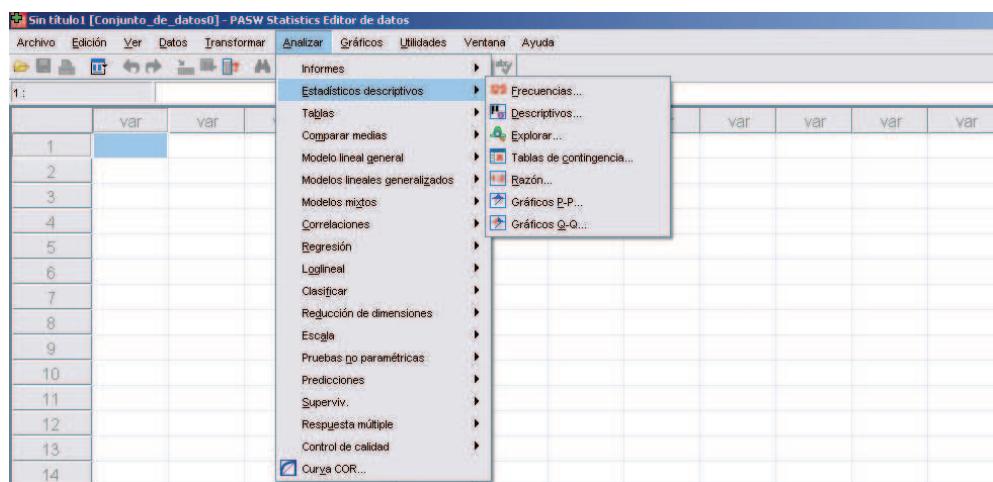
Deskribatu aurreko ariketako aldagaiak, *Vista de variables* leihoa konsultatuz.

2. ESTATISTIKA DESKRIBATZAILEA

2.1. Sarrera

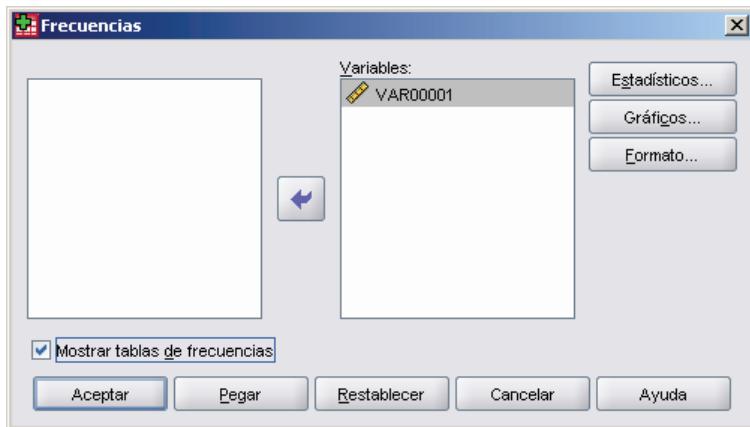
Ingeniaritzan, Biologian, Elikagaien Teknologian, Ekonomian, Medikuntzan edo beste zenbait arloetako kasu praktiko anitzetan, azterketa estatistiko deskribatzalea beharrezkoa da beste teknika estatistiko bat aplikatu baino lehen. Gai honetan azterketa estatistiko deskribatzaleak egiteko SPSS programak eskaintzen dituen zenbait teknika adieraziko dira.

Hasteko, *Analizar > Estadísticos descriptivos* aukeratuz (2.1. irudia), procedura ezberdinak hauta daitezke. Analisi estatistiko deskribatzalea egiteko, gai honetan *Frecuencias, Descriptivos* eta *Explorar* prozedurak azpimarratuko dira.



2.1. irudia.

Frecuencias prozedura aukeratzen bada, aldagai estatistikoaren maiztasun-taula, estatistiko deskribatzailak eta zenbait adierazpen grafiko lortuko dira (2.2. irudia).



2.2. irudia.

Estadísticos botoian klik eginez gero, lau azpikoadrotan sailkatutako estatistiko deskribatzailak azaltzen dira, 2.3. irudian ikus daitekeen modura:



2.3. irudia.

Interesatzen zaizkigun estatistikoen gelaxkak aktibatu ondoren, jarraitu egindo dugu, eta *Gráficos* botoian klik eginez, 2.4. irudiko aukerak azalduko dira:



2.4. irudia.

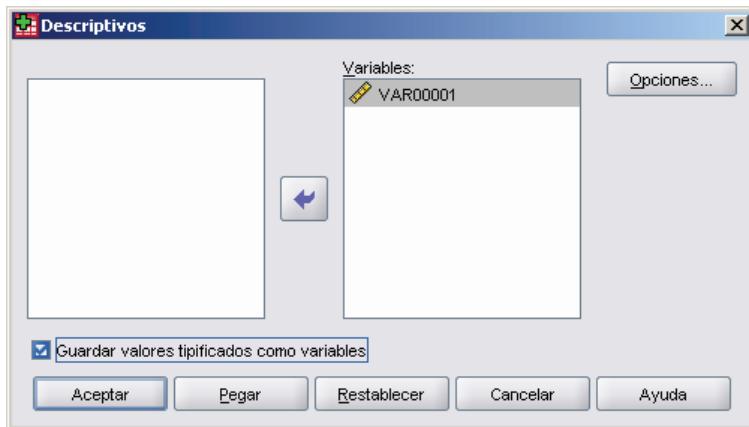
Egiten ari garen azterketa estatistikoan grafikorik ez jartzea erabaki dezakegu, edo barra-grafiko, sektore-diagrama edo histograma (hauetariko bat) eranstea hauta daiteke. Grafikoko balioak maiztasunak edo portzentajeak izan daitezke.

Frecuencias prozeduraz bukatzeko, 2.2. irudiko *Formato* botoian klik eginez emaitzak zein ordenetan adierazi erabaki dezakegu (2.5. irudia).



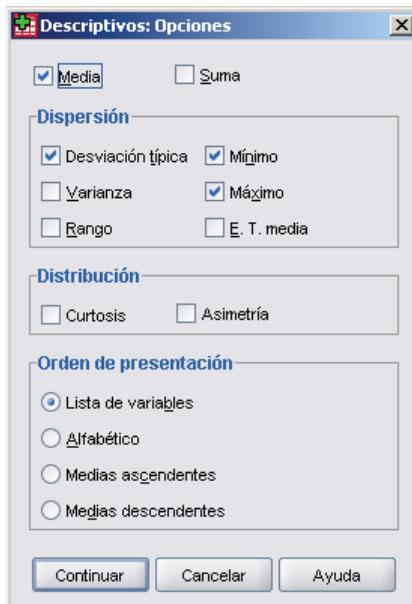
2.5. irudia.

Bestalde, *Descriptivos* prozedura aukeratuz gero, 2.6. irudiko koadroa irekiko da:



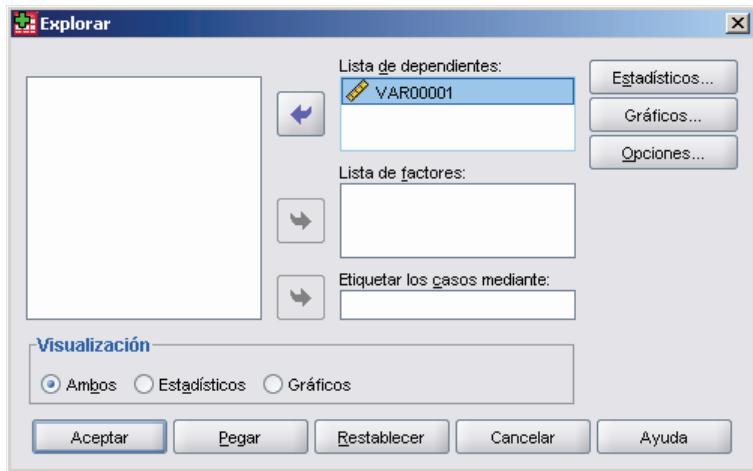
2.6. irudia.

Prozedura honek aldagaiaren balio tipifikatuak gordetzeko aukera ematen du. Kasu horretan, agindua exekutatu ondoren SPSSko datuen leihoa *ZVAR00001* izeneko aldagai berri bat sortuko da. Gainera, *Opciones* botoian klik eginez, estatistiko deskribatzaile batzuk ere kalkula daitezke (2.7. irudia).



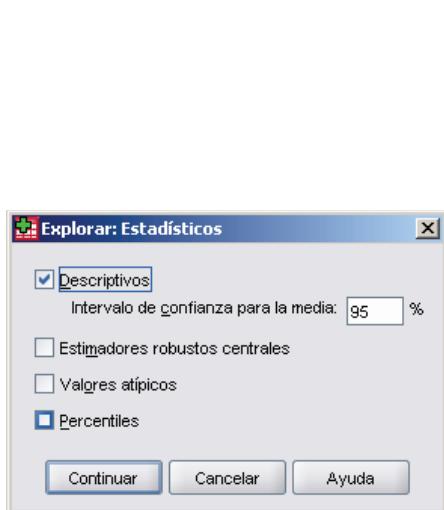
2.7. irudia.

Bukatzeko, *Explorar* prozedura (2.8. irudia) oso interesarria da: balioen baliagarritasuna aztertzeko, balio atipikoak identifikatzeko edo datuen normaltasunaren hipotesia probatzeko (azken hau 6. gaian aztertuko da).

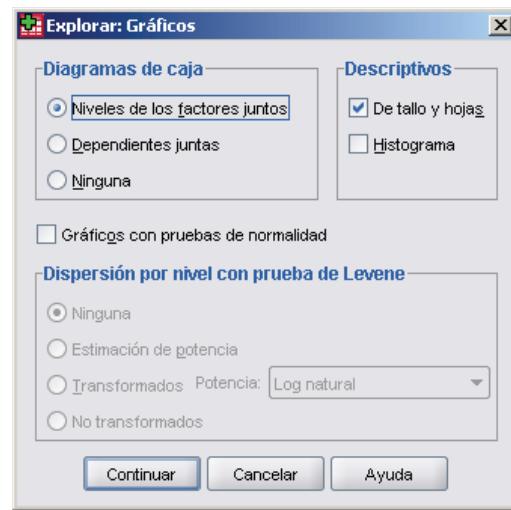


2.8. irudia.

Explorar prozeduraren bidez zenbait estatistiko eta grafiko lor daitezke, 2.9. eta 2.10. irudietan adierazten den moduan.

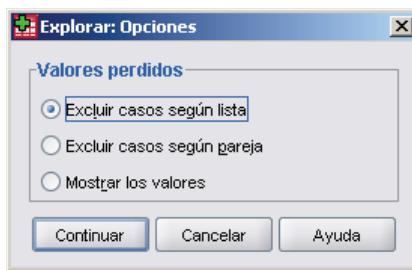


2.9. irudia



2.10. irudia.

Opciones botoian klik eginez, balio galduen buruzko aukera ezberdinak azalduko dira (2.11. irudia).



2.11. irudia.

Prozedura estatistiko hauen funtzionamendua eta erabilpena aztertzeko praktikara pasatuko gara, zenbait adibide ebatziko ditugularik.

2.2. Adibide batzuk

2.2.1. adibidea

Hogeita hamar pieza egiteko elaborazio-denborak (ordutan neurtuta) hauexek izan dira:

128	119	95	97	124	128	142	98	108	120
114	109	124	132	97	138	133	136	120	112
146	128	103	135	114	109	100	111	131	113

- a) Egin elaborazio-denboren analisi estatistiko deskribatzailea.
- b) Adierazi barra-grafikoa eta histograma.
- c) Maiztasun-taulatik ondorioztatu zein den piezen %20k behar izan duen elaborazio-denbora minimoa.
- d) Irudikatu maiztasun metatuen poligonoa. Adierazi bertan zein den P_{80} pertzentila eta azaldu esanahia kontestu honetan.

Ebazpena:

- a) Hasteko, datuak *Denbora* izeneko aldagaien idatzi dira. Gero, *Analizar > Estadísticos descriptivos > Frecuencias* prozedura aukeratu da, eta *Estadísticos* botoian klik egin ondoren minimoa, maximoa, batez bestekoa, mediana, moda, desbideratze tipikoa, kuartilak, 80. ordenako pertzentila, alborapena eta kurtosia hautatu dira. Agindu hau exekutatuz gero, hurrengo emaitzen leihoa (2.12. irudia) irekiko da:

The screenshot shows the SPSS Statistics Visor window. The title bar reads '*Resultado4descrip.svy [Documento3] - PASW Statistics Visor'. The menu bar includes Archivo, Edición, Ver, Datos, Transformar, Insertar, Formato, Analizar, Gráficos, Utilidades, Ventana, Ayuda. The toolbar has various icons for file operations. The left sidebar shows a tree view with 'Resultado', 'Log', 'Frecuencias' (selected), 'Título', 'Notas', 'Conjunto de datos activo', 'Estadísticos', 'Denbora', and 'Gráfico de barras'. The main area contains the following tables:

Estadísticos		
Denbora		
N	Válidos	30
	Perdidos	0
Media		118,80
Mediana		119,50
Moda		128
Desv. típ.		14,490
Asimetría		,006
Curtosis		-,997
Mínimo		95
Máximo		146
Percentiles	25	108,75
	50	118,50
	75	131,25

Denbora					
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado	
Válidos	95	1	3,3	3,3	3,3
	97	2	6,7	6,7	10,0
	98	1	3,3	3,3	13,3
	100	1	3,3	3,3	16,7
	103	1	3,3	3,3	20,0
	108	1	3,3	3,3	23,3
	109	2	6,7	6,7	30,0

2.12. irudia.

Emaitzentz leihotan horretan estatistikoen 2.1. taula adierazten da:

2.1. taula.

Denbora

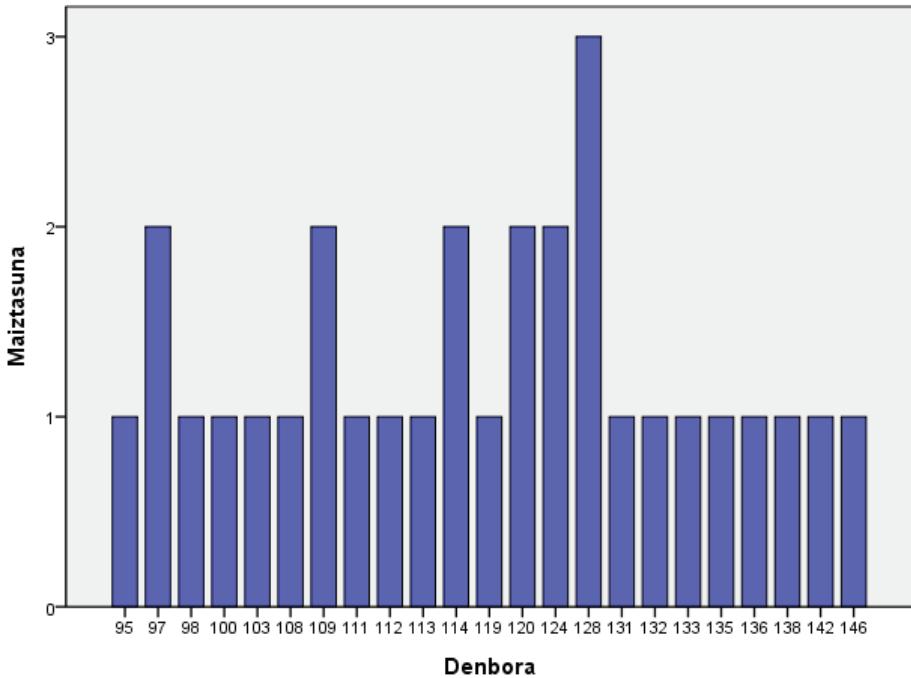
N	Válidos	30
	Perdidos	0
Media		118,80
Mediana		119,50
Moda		128
Desv. Tip.		14,490
Asimetría		,006
Curtosis		-,997
Mínimo		95
Máximo		146
Percentiles	25	108,75
	50	119,50
	75	131,25

Taula honetan adierazten den moduan, adibidean 30 pieza egiteko behar izan diren denborak neurtu dira. Pieza guztien elaborazio-denborak dira ezagunak (baliozkoak 30 eta galduak 0). Piezak egiteko behar izan den denbora minimoa 95 ordukoa izan da eta denbora gehien behar izan duen piezak 146 ordu behar izan ditu. Batez besteko denbora 118,8 ordukoa izan da (hau da, 118,8 ordu erabili izan beharko lirateke bakoitzaz egiteko, pieza guztiek denbora bera behar izango balukete), desbideratze tipikoa 14,49 ordukoa izanik. Moda maiztasun handieneko balioa denez, kasu honetan 128 orduko elaborazio-denbora da gehien errepikatu den denbora.

Posizio-estatistikoen (kuartilen eta pertzentilen) balioek adierazten dutenez, piezen erdiaren elaborazio-denbora [95, 119,5] ordu bitarteko da (mediana edo $P_{50} = 119,5$). Era berean, $P_{25} = 108,75$ orduko pertzentilak adierazten duenez, piezen %25en elaborazio-denbora [95, 108,75] ordu bitartekoa izan da, eta piezen %75en elaborazio-denbora gehienez 131,25 ordukoa izan da ($P_{75} = 131,25$). Bukatzeko, banaketa hau ez da simetrikoa, eskuinerantz alboratua

baizik (alborapena positiboa baita), eta kurtosia negatiboa denez, banaketa platikurtikoa da.

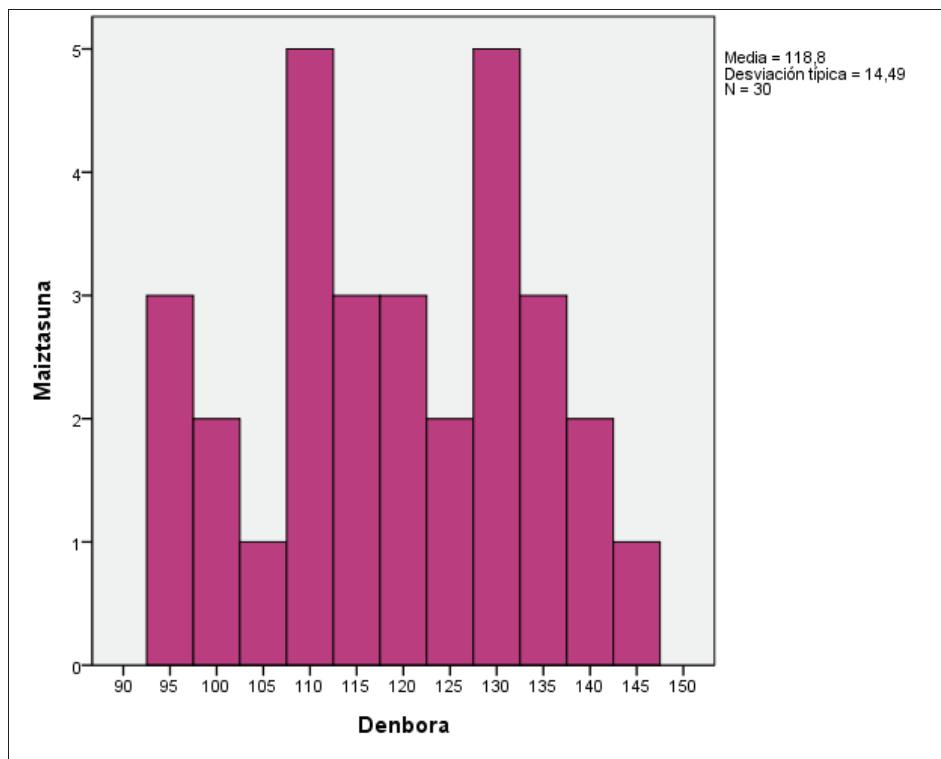
- b) Hona hemen *Denbora* aldagaiari dagokion barra-grafikoa:



2.13. irudia.

Barra-grafikoa lortzeko, *Frecuencias* prozeduran *Gráficos* botoian klik egin eta *Gráfico de barras* (valores *Frecuencias*) aukeratu da. Grafiko honek adierazten duen moduan, denobra minimoa 95 ordu eta maximoa 146 ordu izan dira. Maiztasun handiena izan duen denbora (moda) 128 ordukoa izan da, gainontzeko elaborazio-denborak pieza baterako edo bi piezatarako behar izan direlarik.

Frecuencias prozedurako *Gráficos* aukerak ere badu histograma egiteko gaitasuna. 2.14. irudia *Denbora* aldagaiari dagokion histograma da.



2.14. irudia.

- c) Maiztasun-taularen (2.2. taularen) lehenengo zutabean aldagaiaren baliozko datuak daude. Bigarren zutabean maiztasunak adierazi dira. Hirugarren zutabean balio guztien maiztasunak ehunekotan adierazi dira eta laugarren zutabean baliozko balioen portzentajeak jaso dira. Portzentaje metatuuen zutabean maiztasun erlatibo metatuak ehunekotan adierazi dira. Hona hemen *Denbora* aldagaiari dagokion maiztan-taula:

2.2. taula.

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
95	1	3,3	3,3	3,3
97	2	6,7	6,7	10,0
98	1	3,3	3,3	13,3
100	1	3,3	3,3	16,7
103	1	3,3	3,3	20,0
108	1	3,3	3,3	23,3
109	2	6,7	6,7	30,0
111	1	3,3	3,3	33,3
112	1	3,3	3,3	36,7
113	1	3,3	3,3	40,0
114	2	6,7	6,7	46,7
119	1	3,3	3,3	50,0
120	2	6,7	6,7	56,7
124	2	6,7	6,7	63,3
128	3	10,0	10,0	73,3
131	1	3,3	3,3	76,7
132	1	3,3	3,3	80,0
133	1	3,3	3,3	83,3
135	1	3,3	3,3	86,7
136	1	3,3	3,3	90,0
138	1	3,3	3,3	93,3
142	1	3,3	3,3	96,7
146	1	3,3	3,3	100,0
Total	30	100,0	100,0	

Horrela, piezen %20k behar izan duen elaborazio-denbora minimoa zehazteko 80. ordenako pertzentila zehaztuko da. Maiztasun-taulan portzentaje metatuaren zutabeko 80 zenbakia 132 balioa dagokio, hau da, datuen %80

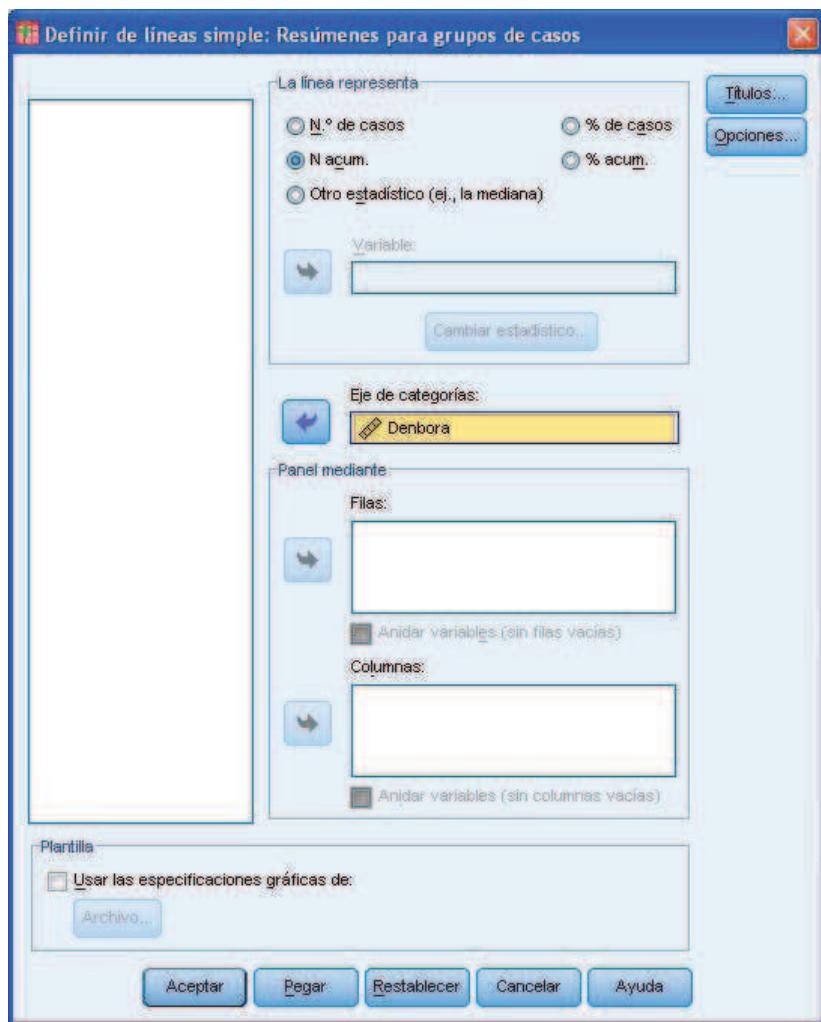
egiteko gehienez 132 ordu behar dira. Ondorioz, piezen %20ren elaborazio-denbora gutxienez 132 ordukoa da.

- d) Maiztasun metatuen poligonoa irudikatzeko, *Gráficos > Cuadros de diálogos antiguos > Líneas* prozeduraren *Gráficos de líneas* azpikoadroan *Resúmenes para grupos de casos* aukeratuko da (2.15. irudia).



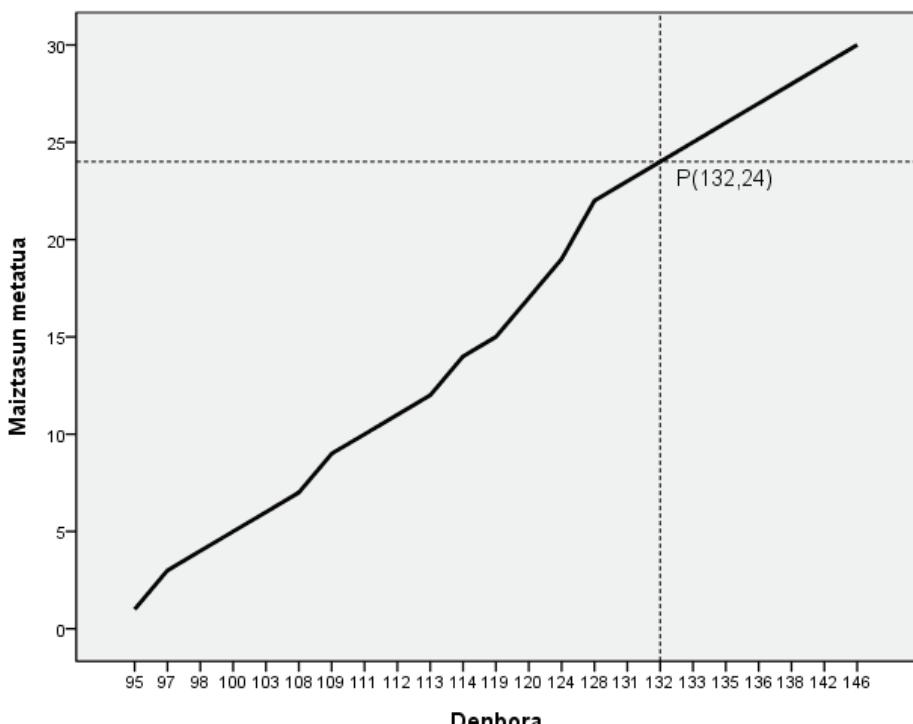
2.15. irudia.

Ondoren, 2.16. irudian adierazten den moduan *N acum.* aktibatuko da eta *Eje de categorías* atalean *Denbora* aldagaia sartuko da.



2.16. irudia.

Ondorioz, 2.17. irudiko maiztasun metatuen poligonoa lortuko da. Poligonoko P puntuaren abzisa-koordenatuaren balioa da 80. ordenako pertzentila, hots, $P_{80} = 132$. Bere esanahia kontestu honetan: 24 piezaren (%80ren) elaborazio-denbora gehienez 132 ordukoa izan da.



2.17. irudia.

2.2.2. adibidea

Hiri bateko A estazioan egun batean jaso diren orduz orduko tenperaturak eta ibilgailu kopurua neurtu dira (2.3. taula).

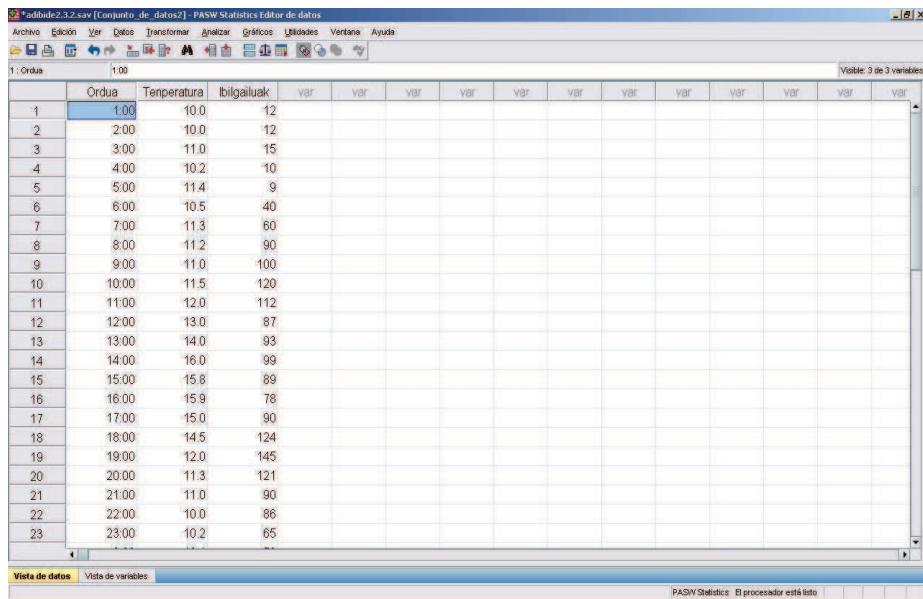
- Adierazi grafikoki tenperaturen bilakaera eta trafikoaren bilakaera.
- Lortu bi aldagaien minimoa, maximoa, batez bestekoa eta desbideratze tipikoa.
- Egia al da arratsaldeko bostetako tenperatura ibilgailu kopurua baino erlatiboki handiagoa dela?

2.3. taula.

Ordua	Tenperatura	Ibilgailu kopurua
01:00	10,0	12
02:00	10,0	12
03:00	11,0	15
04:00	10,2	10
05:00	11,4	9
06:00	10,5	40
07:00	11,3	60
08:00	11,2	90
09:00	11,0	100
10:00	11,5	120
11:00	12,0	112
12:00	13,0	87
13:00	14,0	93
14:00	16,0	99
15:00	15,8	89
16:00	15,9	78
17:00	15,0	90
18:00	14,5	124
19:00	12,0	145
20:00	11,3	121
21:00	11,0	90
22:00	10,0	86
23:00		
00:00	10,1	50

Ebazpena:

- a) Lehenengo *Ordua*, *Tenperatura* eta *Ibilgailuak* izeneko hiru aldagaietan 24 kasuez osaturiko matrizea idatziko da SPSSko datuen leihuan (2.18. irudia).



	Ordua	Tenperatura	Ibilgailuak	var														
1	1:00	10.0	12															
2	2:00	10.0	12															
3	3:00	11.0	15															
4	4:00	10.2	10															
5	5:00	11.4	9															
6	6:00	10.5	40															
7	7:00	11.3	60															
8	8:00	11.2	90															
9	9:00	11.0	100															
10	10:00	11.5	120															
11	11:00	12.0	112															
12	12:00	13.0	87															
13	13:00	14.0	93															
14	14:00	16.0	99															
15	15:00	15.8	89															
16	16:00	15.9	78															
17	17:00	15.0	90															
18	18:00	14.5	124															
19	19:00	12.0	145															
20	20:00	11.3	121															
21	21:00	11.0	90															
22	22:00	10.0	86															
23	23:00	10.2	65															

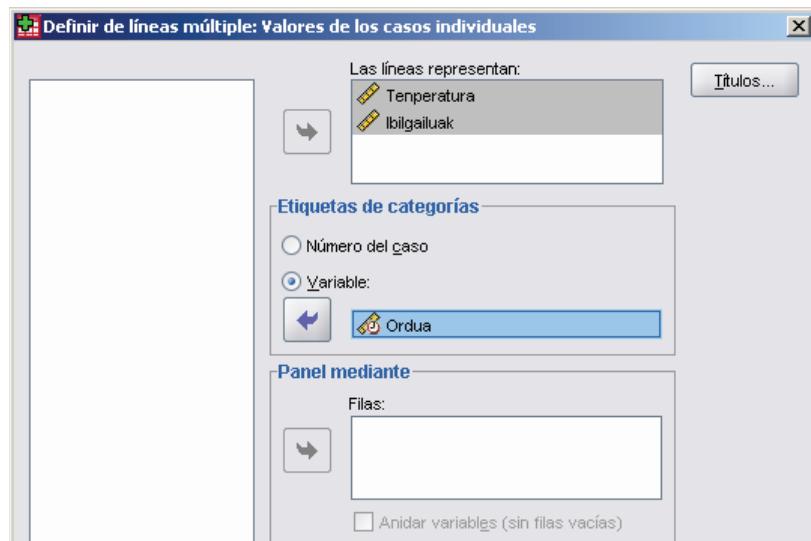
2.18. irudia.

Tenperatura eta *Ibilgailuak* aldagaien bilakaera adierazteko *Graficos* aukeraren barruan, *Líneas* har daiteke. Bertan, *Gráficos de líneas* azpikoadroan *Múltiple* hautatuz, *Valores individuales de los casos* botoia aktibatuko dugu (2.19. irudia).



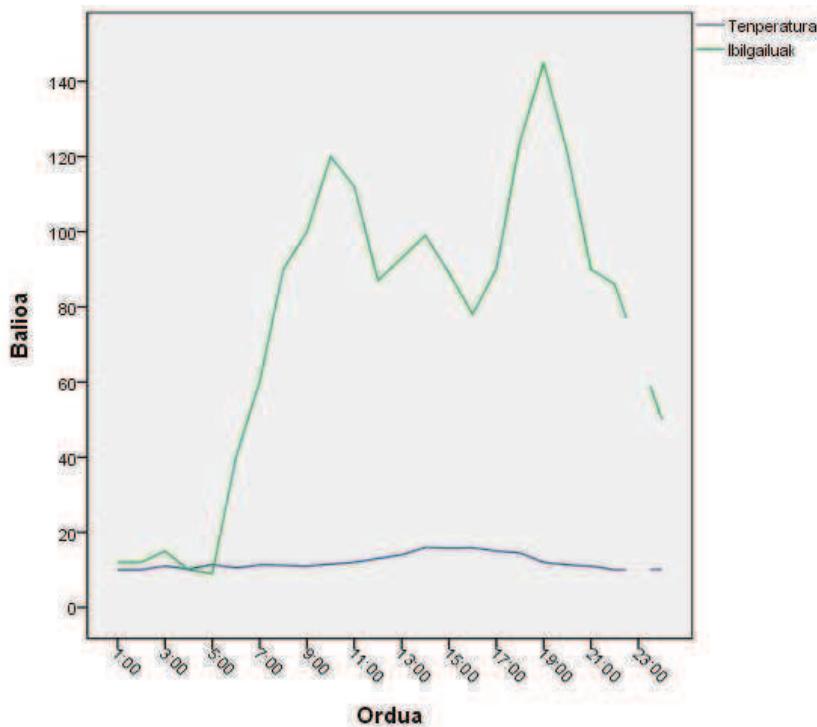
2.19. irudia.

2.20. irudian adierazten den moduan, bi aldagaiak *Las líneas representan* izeneko atalean sartuko dira eta *Etiquetas de variables > Variable* atalean *Ordua* aldagai sartuko da.



2.20. irudia.

Agindu hau exekutatu ondoren, lerro-grafiko bakar batean adieraziko da *Tenperatura* eta *Ibilgailuak* aldagaiiek izan duten bilakaera egun zehatz horretan A estazioan (2.21. irudia).

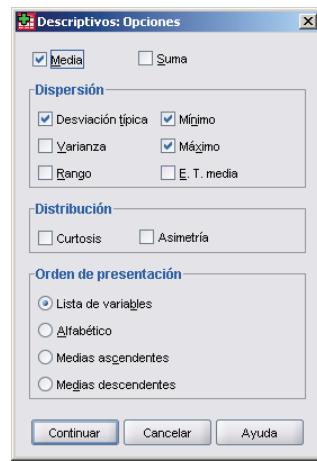


2.21. irudia.

- b) Minimoa, maximoa, batez besteko eta desbideratze tipikoaren balioak lortzeko, *Analizar > Estadísticos descriptivos > Descriptivos* prozedura hauta daiteke. *Descriptivos* koadroan azterketako aldagaiak aukeratuko dira (2.22. irudia) eta *Descriptivos: Opciones* azpikoadroan eskatutako lau estatistiko deskribatzaileen gelaxkak aktibatuko dira (2.23. irudia).



2.22. irudia.



2.23. irudia.

Agindua exekutatu ondoren 2.4. taulako estatistiko deskribatzaileak lortuko dira:

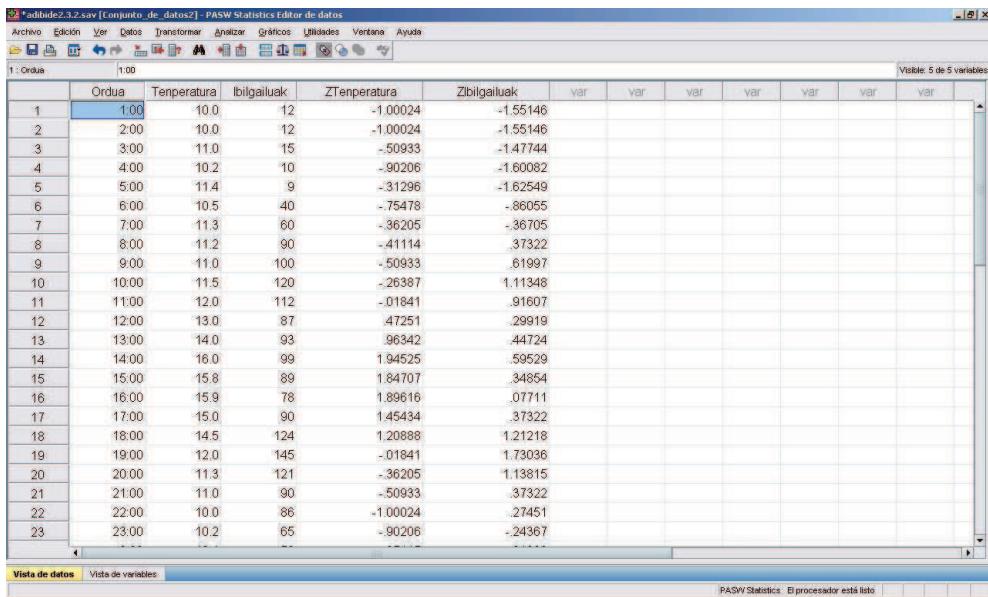
2.4. taula.

	N	Mínimo	Máximo	Media	Desv. tip.
Tenperatura	23	10,0	16,0	12,117	2,0440
Ibilgailuak	23	9	145	75,30	41,381
N válidos	23				

Bai taulak bai lerro-grafikoak balio galdu bat dagoela adierazten dute, hain zuzen ere 23:00etako balio galdua (ikus 2.3. taula).

- c) Atal honen erantzuna lortzeko, 2.22. irudian *Guardar valores tipificados como variables* delakoan klik egingo da eta datuen leihoa bi aldagai berri sortuko dira: *ZTenperatura* eta *ZIbilgailuak* (2.24. irudia). Aldagai hauek *Tenperatura* eta *Ibilgailuak* aldagaien tipifikatuak dira. Arratsaldeko 17:00etako balio tipifikatuek erakusten dutenez, temperaturari dagokion balio tipifikatua (1,454) trafikoari dagokion balio tipifikatua (0,373) baino handiagoa da. Beraz,

arratsaldeko bostetako temperatura ibilgailu kopurua baino erlatiboki handiagoa da.



	Ordua	Temperatura	Ibilgailuak	ZTemperatura	Zibilgailuak	var						
1	1.00	10.0	12	-1.00024	-1.55146							
2	2:00	10.0	12	-1.00024	-1.55146							
3	3:00	11.0	15	-.50933	-.147744							
4	4:00	10.2	10	-.90206	-.160082							
5	5:00	11.4	9	-.31296	-.162549							
6	6:00	10.5	40	-.75478	-.86055							
7	7:00	11.3	60	-.36205	-.36705							
8	8:00	11.2	90	-.41114	.37322							
9	9:00	11.0	100	-.50933	.61997							
10	10:00	11.5	120	-.26387	1.11348							
11	11:00	12.0	112	-.01841	.91607							
12	12:00	13.0	87	.47251	.29919							
13	13:00	14.0	93	.96342	.44724							
14	14:00	16.0	99	1.94525	.59529							
15	15:00	15.8	89	1.84707	.34854							
16	16:00	15.9	78	1.89616	.07711							
17	17:00	15.0	90	1.45434	.37322							
18	18:00	14.5	124	1.20888	1.21218							
19	19:00	12.0	145	-.01841	1.73036							
20	20:00	11.3	121	-.36205	1.13815							
21	21:00	11.0	90	-.50933	.37322							
22	22:00	10.0	86	-.100024	.27451							
23	23:00	10.2	65	-.90206	-.24367							

2.24. irudia.

2.2.3. adibidea

Prezioak.sav izeneko fitxategian (2.25. irudia) lau aldagai daude: *Kodea* izeneko aldagaiak 400 produktuak identifikatzeko erabiltzen da; *Lantegia* izeneko aldagaiak produktu bakoitza zein lantegitan egin den adierazten da; *Prezioa* aldagaiak produktuaren prezioa (eurotan) merkatuan azaltzen du, eta *Kostua* aldagaiak produktu bakoitzen elaborazio-kostua (eurotan) adierazten du.

	Kodea	Lantegia	Kostua	Prezioa	var										
1	1	A	270	569.00											
2	2	C	187	401.00											
3	3	C	120	213.50											
4	4	C	132	218.00											
5	5	C	210	449.00											
6	6	C	135	320.00											
7	7	C	187	359.00											
8	8	C	97	218.00											
9	9	C	127	278.00											
10	10	C	135	239.00											
11	11	C	165	302.00											
12	12	C	120	282.50											
13	13	C	142	276.50											
14	14	C	168	350.00											
15	15	C	135	272.00											
16	16	C	150	407.00											
17	17	C	142	459.00											
18	18	A	275	1036.50											
19	19	C	142	422.00											
20	20	C	115	261.50											
21	21	C	150	387.50											
22	22	C	127	216.50											
23	23	C	111	239.00											

2.25. irudia.

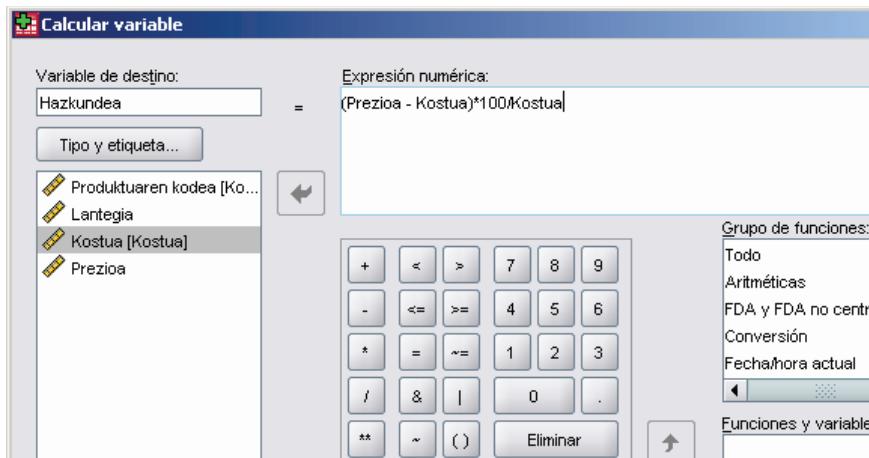
- a) Lortu elaborazio kostuarekiko hazkundea ehunekotan neurzen duen *Hazkundea* izeneko aldagai, hurrengo adierazpena erabiliz:

$$\text{Hazkundea} = (\text{Prezioa} - \text{Kostua}) * 100 / \text{Kostua}$$

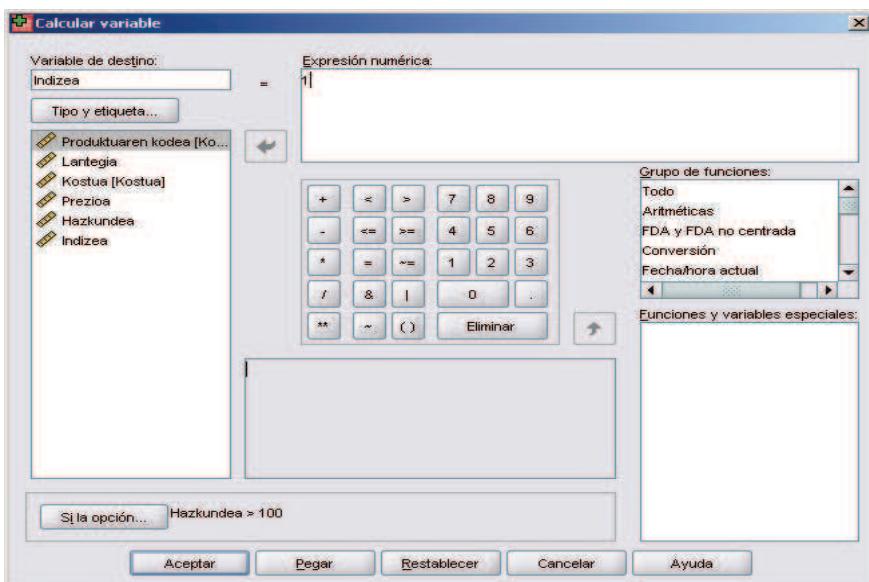
- b) Ondoren, definitu *Indizea* izeneko aldagai berri bat, zeinaren balioak hauek izango diren: 1, *Hazkundea* > 100 denean; 2, *Hazkundea* (50, 100] bitartekoa den kasuetan eta 3 *Hazkundea* txikiago edo berdin 50 denean.
- c) Aukeratu 3 indizedun kasuak eta gorde fitxategi berri batean.
- d) Irudikatu *Indizea* aldagai azaltzen duen sektore-diagrama.
- e) Kutxa-diagramen bidez aztertu grafikoki lantegiaren araberako hazkundea eta aurkitu kasu atipikoak.

Ebazpena:

- a) *Transformar > Calcular variable* prozedura aplikatuko da *Hazkundea* izeneko aldagai berria lortzeko (2.26. irudia) eta era berean jokatuko da *Indizea* aldagai lortzeko (2.27. irudia).



2.26. irudia.



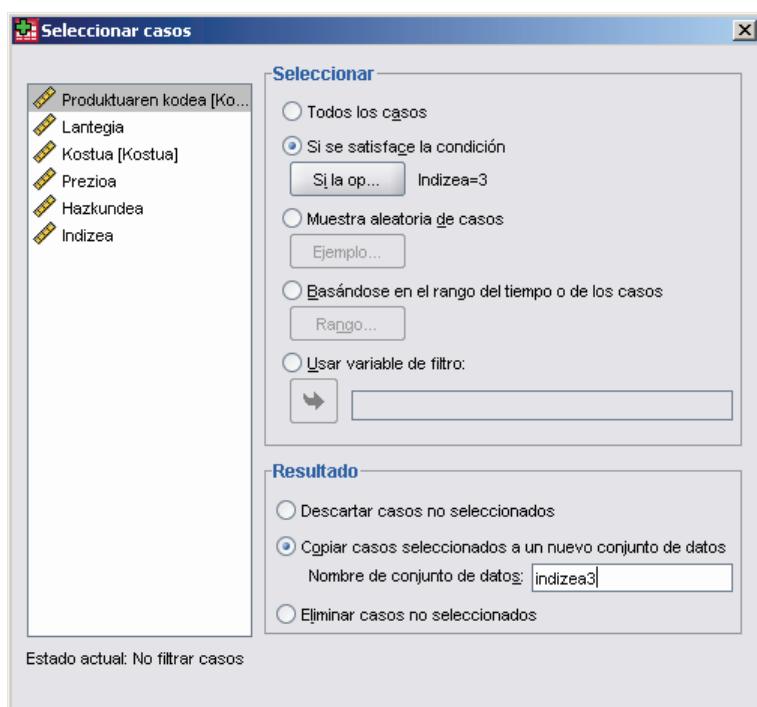
2.27. irudia.

Horrela, a) eta b) ataletan eskatutako bi aldagai berriak lortuko dira (2.28. irudia):

Kodea	Lantegia	Kostua	Prezioa	Hazkundea	Indizea	var							
1	1	A	270	569.00	111.13	1							
2	2	C	187	401.00	114.44	1							
3	3	C	120	213.50	78.66	2							
4	4	C	132	218.00	65.78	2							
5	5	C	210	449.00	114.32	1							
6	6	C	135	320.00	137.92	1							
7	7	C	187	359.00	91.98	2							
8	8	C	97	218.00	124.74	1							
9	9	C	127	278.00	118.90	1							
10	10	C	135	239.00	77.70	2							
11	11	C	165	302.00	83.59	2							
12	12	C	120	282.50	136.40	1							
13	13	C	142	276.50	94.72	2							
14	14	C	168	350.00	108.96	1							
15	15	C	135	272.00	102.23	1							
16	16	C	150	407.00	172.24	1							
17	17	C	142	459.00	223.24	1							
18	18	A	275	1036.50	277.46	1							
19	19	C	142	422.00	197.18	1							
20	20	C	115	261.50	127.39	1							
21	21	C	150	387.50	159.20	1							
22	22	C	127	216.50	70.47	2							
23	23	C	111	239.00	116.29	1							

2.28. irudia.

c) Indizea 3 deneko kasuak aukeratzeko 2.29 irudiko procedura erabili da:



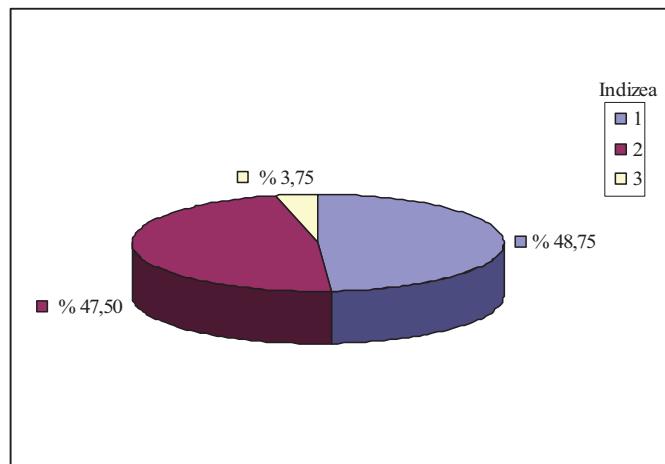
2.29. irudia.

Datos > Seleccionar casos eginez hautatuko dira 3 indizedun kasuak eta datu horiek *indizea3.sav* izeneko fitxategian (2.30. irudian) gordeko dira.

Indizea3.sav [Conjunto_de_datos7] - PASW Statistics Editor de datos															
Archivo Edición Ver Datos Transformer Analizar Gráficos Utilidades Ventana Ayuda															
1: Kodes 136 Visible: 6 de 6 variables															
	Kodes	Lantegia	Kostua	Prezioa	Hazkundea	Indiza	var								
1	136	C	150	221,00	47,83	3									
2	160	A	474	659,00	38,91	3									
3	205	A	525	666,50	27,07	3									
4	217	C	277	345,20	24,62	3									
5	222	C	195	276,50	42,16	3									
6	226	C	157	231,50	47,45	3									
7	230	C	150	224,00	49,83	3									
8	268	C	150	222,50	48,83	3									
9	290	A	362	513,50	41,89	3									
10	302	C	150	222,50	48,83	3									
11	320	C	150	219,50	46,82	3									
12	345	C	165	221,00	34,35	3									
13	372	C	157	212,00	35,03	3									
14	373	C	165	224,00	36,17	3									
15	381	C	180	269,00	49,86	3									
16															
17															
18															
19															
20															
21															
22															
23															

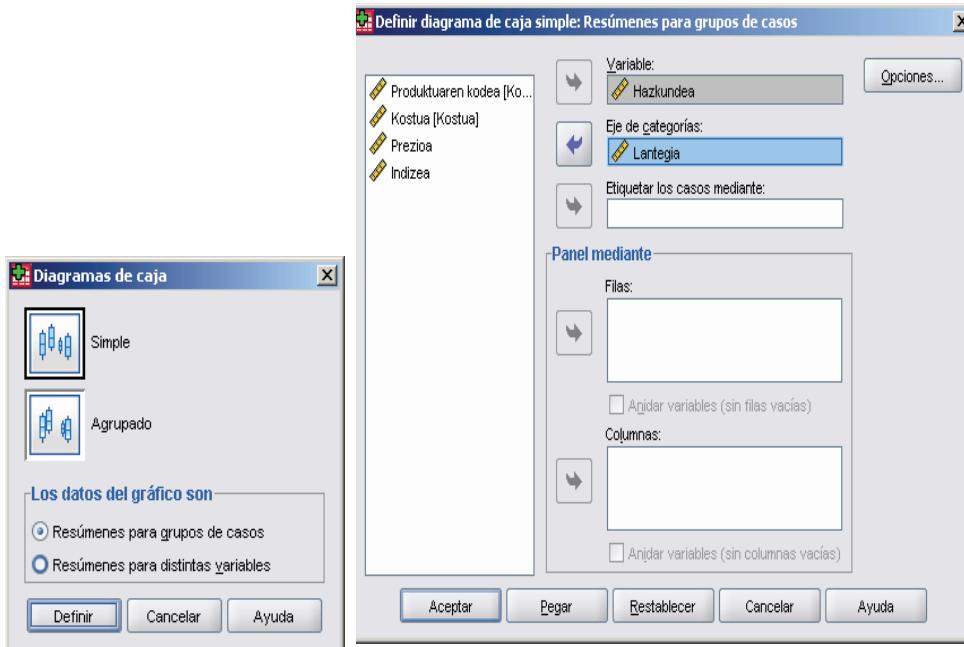
2.30. irudia.

d) Hauxe da *Indiza* aldagaiari dagokion sektore-diagrama (2.31. irudia):



2.31. irudia.

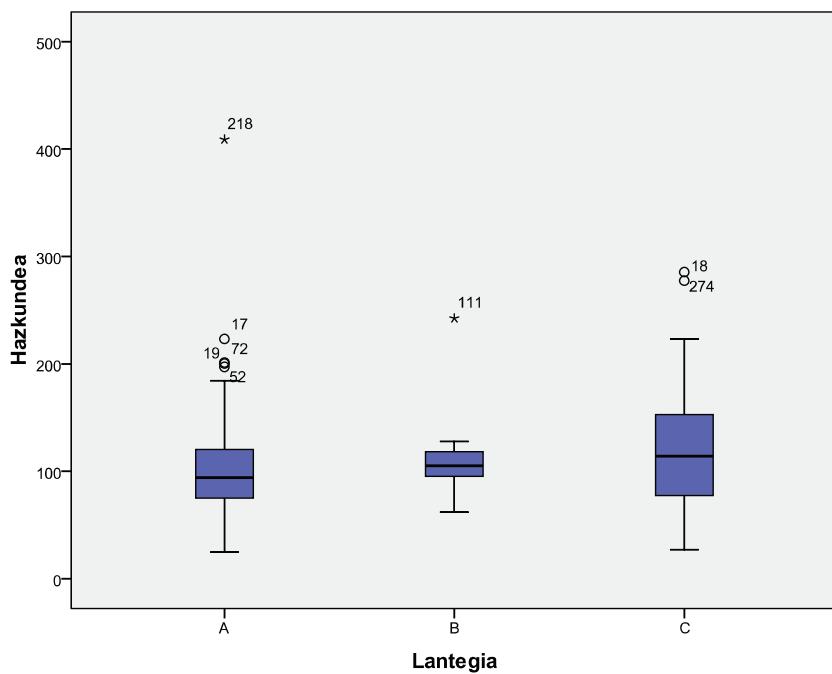
- e) SPSSko menuan *Gráficos > Cuadros de diálogos antiguos > Diagramas de caja* hautatuz 2.32 irudiko koadroa irekiko da eta 2.33. irudian adierazi den moduan eraginez, 2.34. irudiko kutxa-diagrama lortuko da.



2.32. irudia.

2.33. irudia.

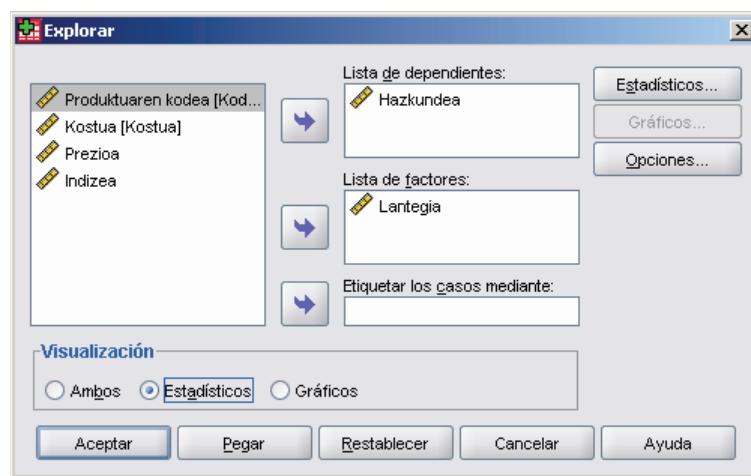
2.34. iruditik ondoriozta daitekeenez, B lantegian egindako produktuen prezioaren hazkundea beste lantegietan baino nabariki txikiagoa izan da. Hazkunde handiena C lantegiko produktuena izan da. Hazkunde handiena (%400 baino handiagoa) A lantegian egindako produktu batek izan du, baina kasu hori muturreko balio bati dagokio. Muturreko balioak (izarrak) batzueta gaizki kalkulatutako/sartutako datuak dira, baina hori ez da oraingoan gertatutakoa. Bi muturreko balio daude: A lantegiko 218 kasua eta B lantegiko 111 kasua.



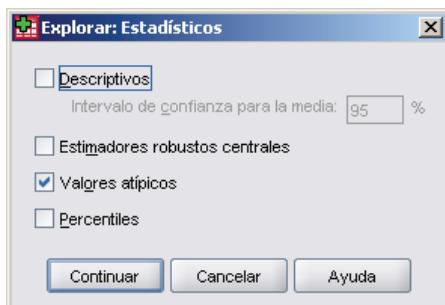
2.34. irudia.

Halaber, 2.34. irudiak kasu atipikoak erakusten ditu: zirkunferentzia txikiak dira eta alboan duten zenbakia kasua da.

Aldagai baten balio atipikoak eta muturrekoak aurkitzeko prozedura *Estadísticos descriptivos > Explorar* da. Aztergai dugun adibidean 2.35. eta 2.36. irudietan adierazten den moduan eragingo dugu.



2.35. irudia.



2.36. irudia.

Ondorioz, 2.5. eta 2.6. tauletan adierazitako irteerak lortuko dira.

2.5. taula.

Resumen del procesamiento de los casos

Lantegia	Casos					
	Válidos		Perdidos		Total	
	N	Porcentaje	N	Porcentaje	N	Porcentaje
Hazkundea A	307	100.0%	0	.0%	307	100.0%
B	25	100.0%	0	.0%	25	100.0%
C	68	100.0%	0	.0%	68	100.0%

2.5. taulan zenbat datu aztertu diren adierazten da, kasu honetan 400 produktutik 307 produktu A lantegikoak, 25 B lantegikoak eta 68 C lantegikoak dira.

Bukatzeko 2.6. taulak jasotzen ditu bost hazkunde txikienak eta bos hazkunde handienak zein produktutan jaso diren. *Valor* zatabean hazkundeen balioak daude eta *Número del caso* zatabean zein produktuk (kasuk) duen ondoko hazkundea adierazten du.

2.6. taula.
Valores extremos

		Nº del caso	Valor	A lantegia					B lantegia					C lantegia				
				Mayores	1	218	408.92		Mayores	1	111	242.46		Mayores	1	274	285.48	
Hazkundea	A lantegia	2	17	223.24					2	45	127.88			2	18	277.46		
		3	52	201.34					3	206	125.08			3	371	223.26		
		4	72	200.28					4	291	123.89			4	106	209.61		
		5	19	197.18					5	303	123.89			5	240	190.24		
		1	217	24.62					1	126	61.87			1	205	27.07		
	Menores	2	345	34.35					2	386	80.89			2	160	38.91		
		3	372	35.03					3	326	87.58			3	290	41.89		
		4	373	36.17					4	385	90.45			4	307	53.82		
		5	222	42.16					5	273	90.45			5	286	56.97		

2.3. Praktikatzeko ariketak

2.3.1. ariketa

Neurgailu batek hezetasun erlatiboko hurrengo datuak jaso ditu:

29	12	34	17	26	28	23	28	29	34
17	23	12	26	34	29	12	34	28	17

- a) Eraiki aldagai honi dagokion maiztasun-taula.
- b) Kalkulatu minimoa, maximoa, batez bestekoa, moda, desbideratze tipikoa, kuartilak, alborapena eta kurtosia. Azaldu esanahia kontestu honetan.
- c) Eraiki barra-diagrama, maiztasun metatuen poligonoa eta kutxa-diagrama.

2.3.2. ariketa

Hona hemen marka ezagun bateko 20 kotxeren gasolina-kontsumoa (litro) 100 km-ko ibilbidean:

8,1	6,3	4,4	7,0	5,50	6,0	6,7	7,6	8,8	9,3
7,7	7,8	8,8	6,9	7,3	6,8	6,8	8,0	7,7	7,4

- a) Sartu datu hauek aurreko ariketakoz datuen fitxategi berean.
- b) Har itzazu 10. lerroko hezetasuna eta gasolina-kontsumoa. Lortu balio tipifikatuak eta ondorioztatu zein balio den erlatiboki handiagoa.
- c) Egia al da gasolina-kontsumoko datuak hezetasuneko datuak baino kontzentratuago daudela batez bestekoarekiko?
- d) Egin gasolina-kontsumoaren analisi estatistiko deskribatzalea eta interpretatu lortutako emaitzak kontestu honetan.

2.3.3. ariketa

Airearen kalitatea neurten den estazio batean 24 egunetako batez besteko ozono mailak (μgm^{-3}) hauek izan dira:

97	98	99	101	102	104	105	106	107
108	109	-12	111	112	113	23	115	116
117	118	190	120	123	110			

- a) Aurkitu balio atipikoak eta adierazi grafikoki. Ba al dago ezabatu beharreko balioren bat?
- b) Kalkula ezazu datu hauen batez bestekoa eta konpara ezazu balio atipikoak ezabatuta kalkulatuko den batez bestekoarekiko.

3. PROBABILITATE-BANAKETA DISKRETUAK

3.1. Aldagai aleatorio diskretuaren probabilitate-funtzioa eta banaketa-funtzioaren balioak SPSS erabiliz

SPSS programaren bidez banaketa diskretuei dagozkien probabilitate-funtzioa eta banaketa-funtzioaren balioak kalkula daitezke. Horrela, **probabilitate-funtzioaren balioak** lortzeko SPSS softwareak **PDF** (*probability density function*) funtzioa erabiltzen du. Halaber, **banaketa-funtzioaren balioak** lortzeko SPSSren bidez **CDF** (*cumulative distribution function*) funtzioa erabili behar da.

Balioak lortzeko *Transformar > Calcular variable* prozedura aukeratuko da. Irekiko den koadroan *Variable de destino* dioen lekuan lortu nahi den balioa erakutsiko duen aldagaiaren izena idatziko da eta *Expresión numérica* izenaren azpian dagokion funtzioaren adierazpena zehaztuko da. SPSSren laguntzak funtzio hauen zerrenda luzea du.

3.1. taulan gai honetan aztertuko diren probabilitate-funtzioak eta banaketa-funtzioak jaso dira:

3.1. taula.

<i>Sintaxia</i>	<i>Probabilitatea</i>	<i>Zein banaketarena</i>
<i>PDF.BINOM(x, n, p)</i>	$P(X = x)$	<i>Binom(n, p)</i>
<i>PDF.POISSON(x, λ)</i>		<i>Poisson (λ)</i>
<i>PDF.HYPER(x, N, n, a)</i>		<i>Hiper(N, n, a/N)</i>
<i>CDF.BINOM(x, n, p)</i>	$P(X \leq x)$	<i>Binom(n, p)</i>
<i>CDF.POISSON(x, λ)</i>		<i>Poisson (λ)</i>
<i>CDF.HYPER(x, N, n, a)</i>		<i>Hiper(N, n, a/N)</i>

SPSS programak duen kalkulu ahalmenari esker, ez da beharrezkoa konbergentzia-teoremen aplikazioa.

3.2. Adibide batzuk

3.2.1. adibidea

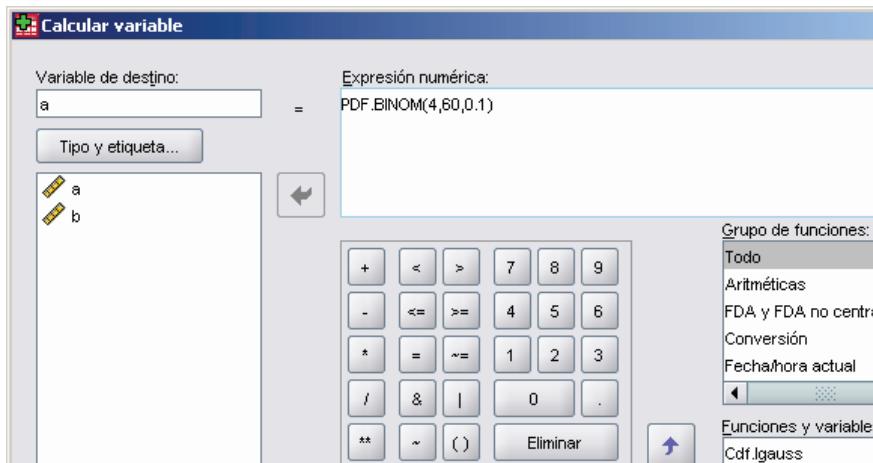
Fabrika batean makinen %10 A motakoa da. Aleatorioki 60 makina hartu dira.

- a) Kalkulatu A motako lau makina izateko probabilitatea.
- b) Lortu A motako gutxienez 5 makina izateko probabilitatea.
- c) A motako zenbat makina itxaron daitezke?

Ebazpena:

Biz $X = \text{"A motako makina kopurua, 60 makinatik"}$ aldagai $n = 60$ eta $p = 0,1$ parametroetako banaketa binomialekoa, $p = P(\text{A motakoa}) = 0,1$ eta $q = P(\text{beste mota bateko}) = 0,9$ balioak edozein makinatarako konstanteak izanik.

- a) 3.1. irudian $P(X = 4)$ balioaren kalkulurako erabiliko den sintaxia adierazten da:



3.1. irudia.

Beraz, A motako lau makina izateko probabilitatea hauxe da:

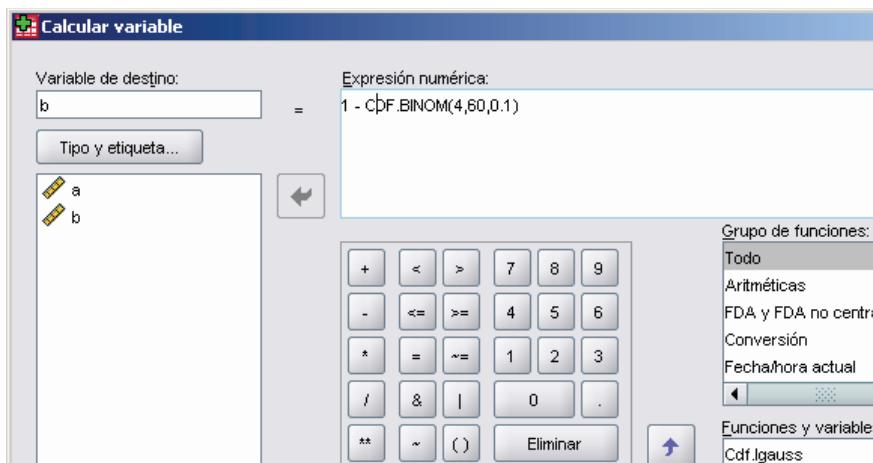
$$P(X=4) = PDF.BINOM(4, 60, 0.1) = 0,1336$$

3.2. irudian jaso da emaitza:

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
	a	b	var	var	var	var	var	v
1	.1336							
2								
3								
4								

3.2. irudia.

- b) Dagokion banaketa-funtzioa erabiliz kalkulatuko da $P(X \geq 5)$ balioa (3.3 irudia eta 3.4. irudia).



3.3. irudia.

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
	Archivo	Edición	Ver	Datos	Transformar	Analizar	Gráficos	Utilidades
1: b		.729041741874539						
		a	b	var	var	var	var	v
1		.1336	.7290					
2								
3								
4								

3.4. irudia.

Ondorioz, gutxienez bost makina izateko probabilitatea hurrengoa da:

$$P(X \geq 5) = 1 - CDF.BINOM(4, 60, 0.1) = 0,7290$$

- c) Banaketa binomialaren batez besteko balioa kalkulatz, A motako

$$E(X) = np = 60 \cdot 0,1 = 6$$

makina itxaron daitezke.

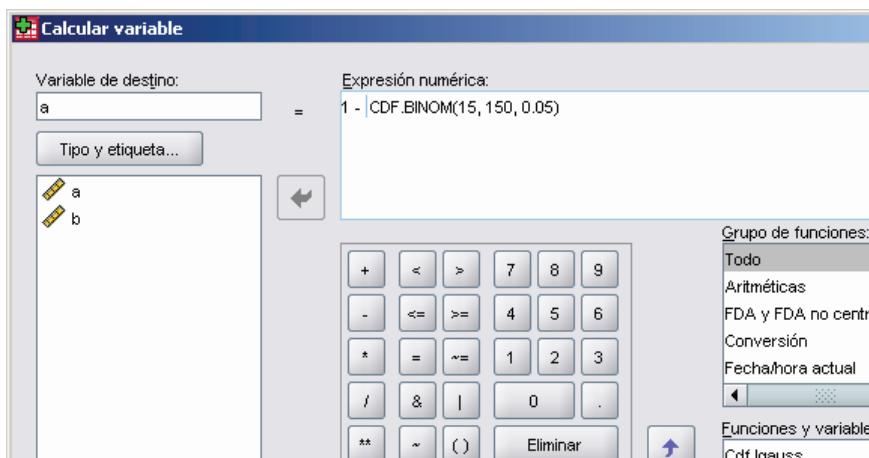
3.2.2. adibidea

Enpresa batek egiten dituen burdinazko habeen %5 akastuna da. Kalitatea kontrolatzeko asmoz, egunero aleatorioki burdinazko 150 habe hartzen dira, eta hauetako 15 habe baino gehiago akastunak badira, ekoizpena gelditzen da. Zein da aleatorioki hartutako egun batean ekoizpena gelditzeko probabilitatea?

Ebazpena:

Biz X = “burdinazko habe akastunen kopurua, burdinazko 150 habetik” aldagai aleatorioa. Edozein habe akastuna izateko probabilitatea 0,05 bada, orduan X aldagaiak $n = 150$ eta $p = 0,05$ parametroetako banaketa binomiala duela onar daiteke.

$P(X > 15)$ probabilitatea kalkulatuz ebatzikoa da egindako galdera (3.5. irudia).



3.5. irudia.

Beraz, aleatorioki hartutako edozein egunetan ekoizpena gelditzeko probabilitatea hauxe da:

$$P(X > 15) = 1 - CDF.BINOM(15, 150, 0.05) = 0,0036$$

3.6 irudian jaso da emaitza:

	a	b	var	var	var	var	var	v
1		.0036						
2								
3								
4								

3.6. irudia.

3.2.3. adibidea

Elikagai enpresa batek egunero, batez beste, sei eskaera jasotzen ditu.

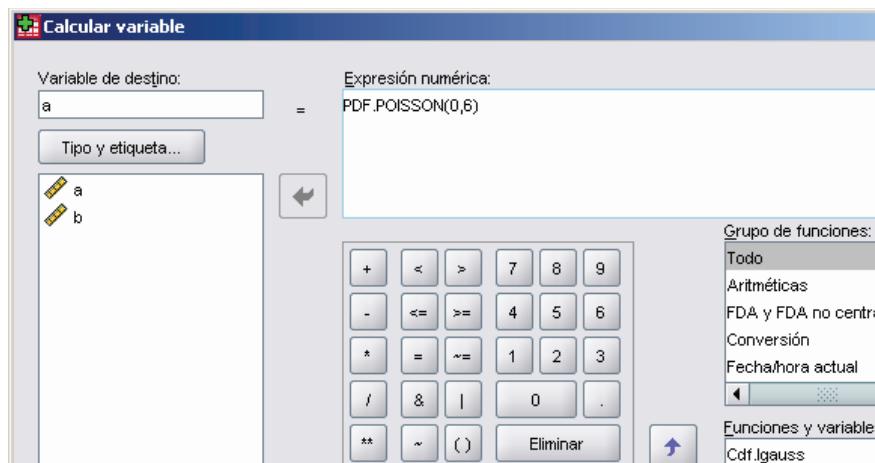
- a) Aleatorioki hartutako egun batean, kalkulatu eskaerarik ez jasotzeko probabilitatea.
- b) Aleatorioki hartutako egun batean, lortu gutxienez eskaera bat eta gehienez lau eskaera jasotzeko probabilitatea.
- c) Aleatorioki hartutako bi egunetan, zein da gehienez 10 eskaera jasotzeko probabilitatea?

Ebazpena:

Demagun X = “elikagai enpresak eguneko jasotzen duen eskaera kopurua” aldagai aleatorioak $\lambda_x = 6$ parametroko Poisson-en banaketa duela.

- a) Atal honen erantzuna 3.7. eta 3.8. irudietan jasota dago:

$$P(X = 0) = PDF.POISSON(0, 6) = 0,0025$$



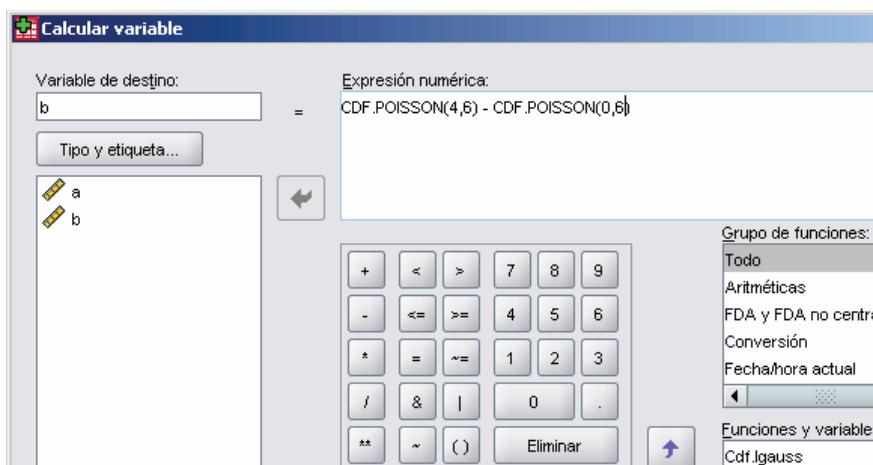
3.7. irudia.

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos	
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda	
1: a	.002478752176666
1	.0025
2	
3	
4	

3.8. irudia.

b) 3.9. eta 3.10. irudietan adierazten da erantzuna:

$$P(1 \leq X \leq 4) = CDF.POISSON(4, 6) - CDF.POISSON(0, 6) = 0,2826$$

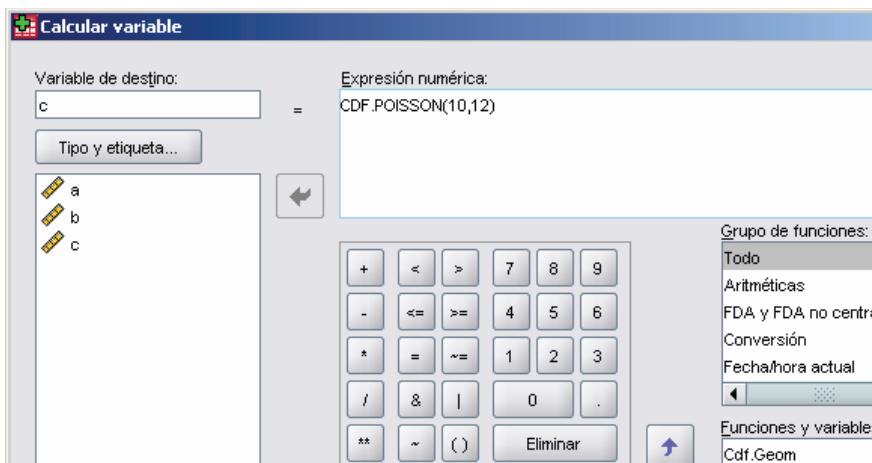


3.9. irudia.

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
	Archivo	Edición	Ver	Datos	Transformar	Analizar	Gráficos	Utilidades
1: b								
1	.0025	2826		var	var	var	var	v
2								
3								
4								

3.10. irudia.

- c) Biz $Y = \text{"elikagai enpresak bi egunetan jasotzen duen eskaera kopurua"}$ aldagai aleatorioia $\lambda_y = 12$ parametroko Poisson-en banaketakoa. Orduan, bi egunetan gehienez 10 eskaera jasotzeko probabilitatea (3.12. irudikoa) 3.11. irudian adierazi den moduan kalkula daiteke:



3.11. irudia.

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos									
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda									
1: c .347229417554172									
	a	b	c	var	var	var	var	v	
1	.0025	.2826	.3472						
2									
3									
4									

3.12. irudia.

$$P(Y \leq 10) = CDF.POISSON(10, 12) = 0,3472$$

3.2.4. adibidea

Araztegi batean dauden 300 hondakinetik 30 hondakinek substantzia ezezagun bat dute. Aleatorioki araztegiko 25 hondakin hartu dira.

- a) Zein da hondakinen %10ek baino gutxiagok substantzia ezezaguna edukitzeko probabilitatea?
- b) Kalkulatu substantzia ezezaguna duten hondakinen kantitatea gutxienez %20 eta gehienez %40 izateko probabilitatea.

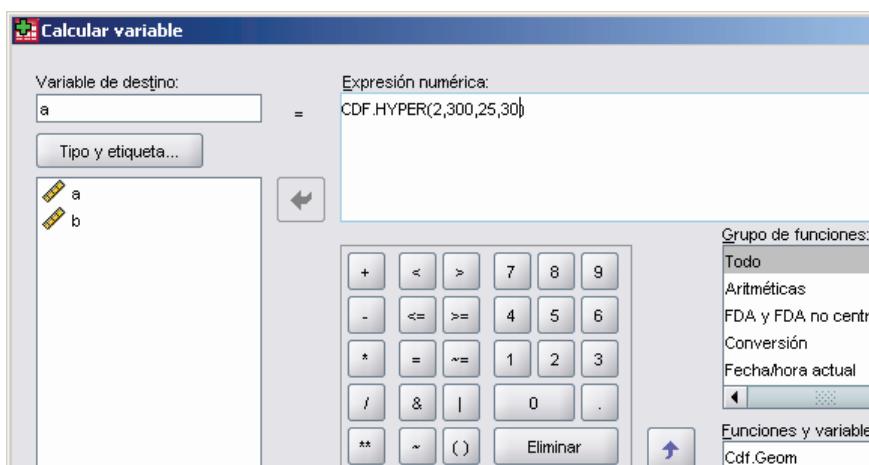
Ebazpena:

Biz $X = \text{"substantzia ezezaguna duen hondakin kopurua, 25 hondakinetik"}$ aldagai aleatorioa. Demagun aldagai honek $N = 300$, $n = 25$ eta $a = 30$ parametroetako banaketa hipergeometrikoa duela.

- a) 3.13. irudian adierazten den moduan,

$$P(X < 2,5) = P(X \leq 2) = \text{CDF.HYPER}(2, 300, 25, 30)$$

kalkulatuko da.



3.13. irudia.

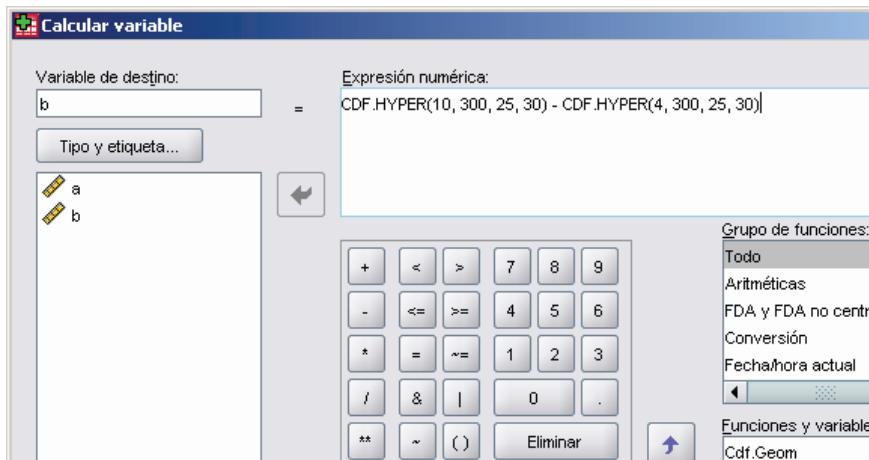
Emaitza 3.14. irudia jaso da:

$$P(X \leq 2) = 0,5324$$

	a	b	var	var	var	var	var	v
1	5324	0.889						
2								
3								

3.14. irudia.

- b) Atal honetan $P(5 \leq X \leq 10)$ probabilitatearen balioa lortu behar da (ikus 3.15. irudia).



3.15. irudia.

$$P(5 \leq X \leq 10) = CDF.HYPER(10, 300, 25, 30) - CDF.HYPER(4, 300, 25, 30)$$

3.16 irudian adierazi da emaitza:

	a	b	var	var	var	var	var	v
1	5324	0.889						
2								
3								

3.16. irudia.

$$P(5 \leq X \leq 10) = 0,0889$$

3.2.5. adibidea

Aholkularitza enpresa batek urtean, batez beste, 1.200 pertsonari ematen die aholkua. Aleatorioki hartutako urte batean, kalkulatu:

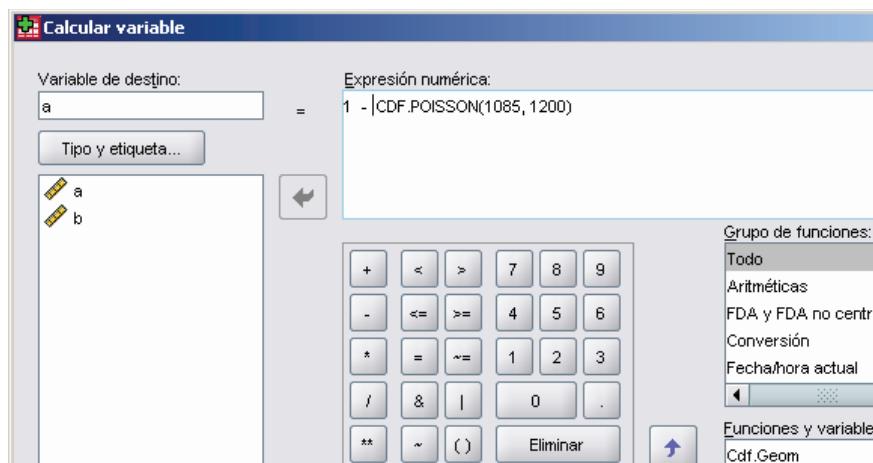
- a) 1.085 pertsona baino gehiagori aholkua emateko probabilitatea.
- b) 1.200-1.300 pertsonari aholkua emateko probabilitatea.

Ebazpena:

Demagun X = “urtean aholkua jasotzen duen pertsona kopurua” aldagai aleatorioak $\lambda = 1.200$ parametroko Poisson-en banaketa duela.

- a) Ebazteko erabilitako sintaxia eta prozedura 3.17. eta 3.18. irudietan dago:

$$P(X > 1085) = 1 - CDF.POISSON(1085, 1200) = 0,9996$$

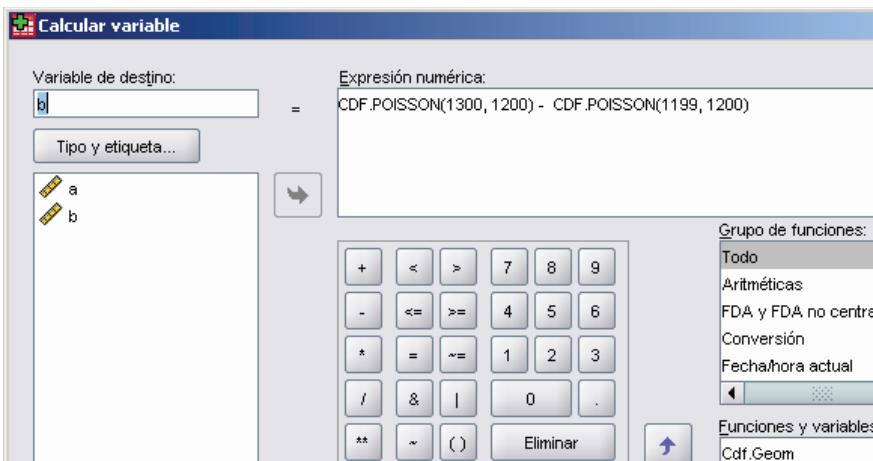


3.17. irudia.

	a	b	var	var	var	var	var	v
1	9996							
2								
3								

3.18. irudia.

b) $P(1.200 \leq X \leq 1.300)$ balioa kalkulatu behar da (ikus 3.19. eta 3.20 irudiak).



3.19. irudia.

	a	b	var	var	var	var	var	v
1	9996	5018						
2								
3								

3.20. irudia.

$P(1.200 \leq X \leq 1.300) = 0,5018$ balioa lortzeko erabili den sintaxia hau da:

$$CDF.POISSON(1300, 1200) - CDF.POISSON(1199, 1200)$$

3.2.6. adibidea

Ordenagailu-areto batean 750 ordenagailu daude eta 450 ordenagailu A sistema eragilea eta 300 ordenagailuk B sistema eragilea erabiltzen dituzte. Ordenagailu bakoitzak sistema eragile bakarra erabiltzen du. Aleatorioki 20 ordenagailu hartu dira.

- a) Zein da 6 ordenagailuk B sistema eragilea erabiltzeko probabilitatea?
- b) Kalkula bedi gutxienez 12 ordenagailuk B sistema eragilea erabiltzeko probabilitatea.
- c) Kalkulatu gutxienez 6 ordenagailuk eta gehienez 9 ordenagailuk A sistema eragilea erabiltzeko probabilitatea.

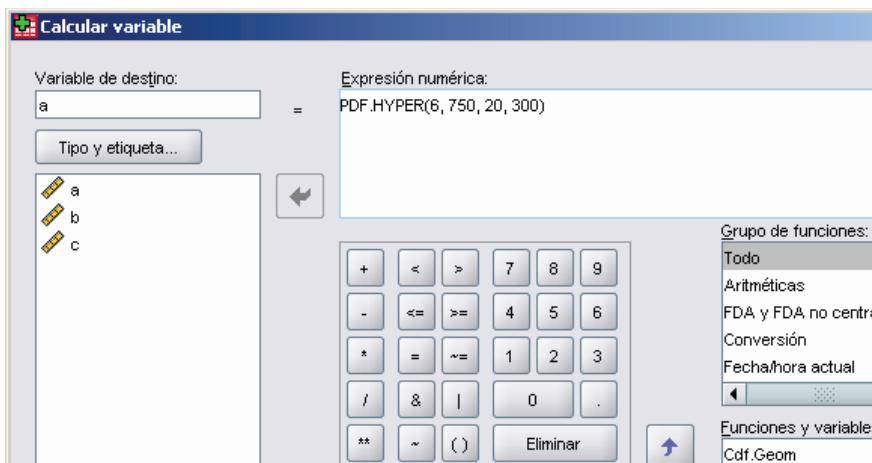
Ebazpena:

Biz X = “hogei ordenagailutik, B sistema eragilea erabiltzen duen ordenagailu kopurua”. X aldagai aleatorioak $N = 750$, $n = 20$ eta $\alpha = 300$ parametroetako banaketa hipergeometrikoa du.

- a) Eskatutako probabilitatearen balioa kalkulatzeko erabiliko den sintaxia hauxe da:

$$P(X=6) = PDF.HYPER(6, 750, 20, 300) = 0,1245$$

Prozedura eta emaitza 3.21. eta 3.22 irudietan jasota daude:



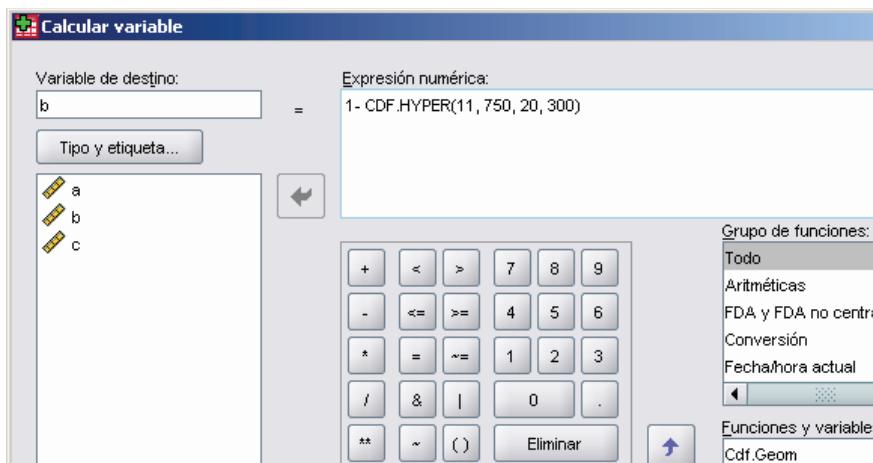
3.21. irudia.

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
	Archivo	Edición	Ver	Datos	Transformar	Analizar	Gráficos	Utilidades
1: a								
1								
2								
3								
4								

3.22. irudia.

- b) $P(X \geq 12)$ probabilitatearen balioa kalkulatzeko urratsak eta balioa bera 3.23. eta 3.24 irudietan daude, erabilitako sintaxia eta emaitza hurrengoak izan direlarik:

$$P(X \geq 12) = 1 - P(X \leq 11) = 1 - CDF.HYPER(11, 750, 20, 300) = 0,0541$$

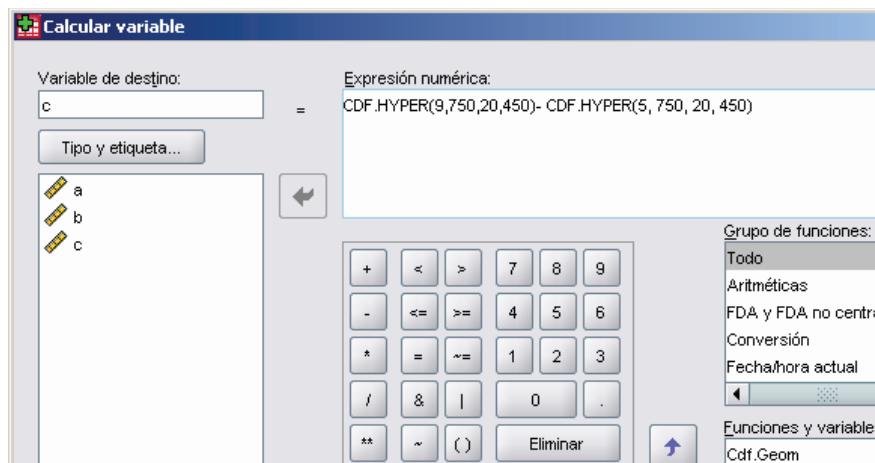


3.23. irudia.

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
Archivo	Edición	Ver	Datos	Transformar	Analizar	Gráficos	Utilidades	Ayuda
1 : b		054102606715177						
			a	b	c	var	var	var
1		.1245		.0541		.		
2								
3								
4								

3.24. irudia.

- c) Biz $Y =$ “hogei ordenagailutik, A sistema eragilea erabiltzen duen ordenagailu kopurua” aldagai aleatorioa. Demagun Y aldagaiak $N_y = 750$, $n_y = 20$ eta $a_y = 450$ parametroetako banaketa duela. Kasu honetan 3.25. eta 3.26. irudietan jaso da ebazpena.



3.25. irudia.

The screenshot shows the PASW Statistics Editor de datos interface. The title bar reads "AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos". The menu bar includes Archivo, Edición, Ver, Datos, Transformar, Analizar, Gráficos, Utilidades, Ventana, and Ayuda. Below the menu is a toolbar with icons for file operations like Open, Save, Print, and Data Manipulation. The main workspace shows a single row of data labeled "1: c" with values ".12294672892197" across eight columns labeled "a", "b", "c", and "var".

3.26. irudia.

Hona hemen erabili den sintaxia eta soluzioa:

$$P(6 \leq Y \leq 9) = CDF.HYPER(9, 750, 20, 450) - CDF.HYPER(5, 750, 20, 450) \\ = 0,1229.$$

3.3. Praktikatzeko ariketak

3.3.1. ariketa

Lanpara fluoreszenteen %87k gutxienez 10.000 orduko iraupena du. Aleatorioki mota horretako 30 lanpara hartu dira.

- a) Zein da 30 lanparatik 26 lanparak gutxienez 10.000 orduko iraupena izateko probabilitatea.
- b) Kalkula bedi 30 lanparatik gutxienez 22 lanparak gutxienez 10.000 orduko iraupena izateko probabilitatea.
- c) Hogeita hamar lanparatik 27 lanpara gutxienez 10.000 orduko iraupenekoak dira. Hogeita hamar lanparetatik aleatorioki hamalau lanpara hartu dira. Zein da hamar lanpara gutxienez 10.000 orduko iraupenekoak izateko probabilitatea?

3.3.2. ariketa

Ingeniaritza-enpresa batek urtean, batez beste, 1.500 lan bukatzen ditu.

- a) Kalkula bedi aleatorioki hartutako urte batean gutxienez 1.200 lan eta gehienez 1.600 lan bukatzeko probabilitatea.
- b) Zein da aleatorioki hartutako hile batean gutxienez 110 lan bukatzeko probabilitatea?

3.3.3. ariketa

Trokelatzeko makina batek 500 pieza egin ditu, zeintzuetatik 450 pieza esportatzekoak diren. 500 piezetatik aleatorioki 200 pieza hartu dira.

- a) Zein da pieza hauetatik guztiak esportatzekoak izateko probabilitatea?
- b) Zein da pieza hauetatik gutxienez 150 eta gehienez 160 pieza esportatzekoak izateko probabilitatea?

3.3.4. ariketa

Bira X : $\text{Binom}(200, 0,01)$ eta Y : $\text{Poisson}(2)$. Kalkulatu $P(X \leq 5)$ eta $P(Y \leq 5)$ eta kalkulatu diferentzia.

4. PROBABILITATE-BANAKETA JARRAITUAK

4.1. Aldagai aleatorio jarraituaren dentsitate-funtzioa, banaketa-funtzioa eta alderantzizko banaketa-funtzioa SPSS erabiliz

Aurreko gaian azaldu den moduan, banaketa jarraituen kasuan ere SPSS programaren bidez dentsitate-funtzioa eta banaketa-funtzioaren balioak kalkula daitezke. Horrela, **dentsitate-funtzioaren balioak** lortzeko SPSS softwareak **PDF** (*probability density function*) funtzioa erabiliko du eta **CDF** (*cumulative distribution function*) funtzioaren bidez **banaketa-funtzioaren balioak** lortuko ditu. Gainera, banaketa jarraituen kasuan SPSS programak **alderantzizko banaketa-funtzioaren balioak** lortzeko aukera ematen du **IDF** (*inverse distribution function*) funtzioaren bidez. *Transformar > Calcular variable* prozedura aukeratu ondoren irekiko den koadroko zenbakizko adierazpenean idatziko dira funtzi hauek.

4.1. taulan gai honetan aztertuko diren **PDF**, **CDF** eta **IDF** funtziok jaso dira:

4.1. taula.

<i>Sintaxia</i>	<i>Balioa</i>	<i>Zein banaketarena</i>
<i>PDF.UNIFORM(x, a, b)</i>	Dentsitate-funtzioa	<i>Uniform(a, b)</i>
<i>PDF.EXP(x, 1/β)</i>		<i>Exp(β)</i>
<i>PDF.NORMAL(x, μ, σ)</i>		<i>N(μ, σ)</i>
<i>PDF.CHISQ(x, gl)</i>		<i>gl</i> askatasun-graduko khi karratu banaketa
<i>PDF.T(x, gl)</i>		<i>gl</i> askatasun-graduko Student-en banaketa
<i>PDF.F(x, gl₁, gl₂)</i>		<i>gl₁</i> eta <i>gl₂</i> askatasun- gradutako <i>F</i> banaketa
<i>CDF.UNIFORM(x, a, b)</i>	<i>P(X ≤ x)</i>	<i>Uniform(a, b)</i>
<i>CDF.EXP(x, 1/β)</i>		<i>Exp(β)</i>
<i>CDF.NORMAL(x, μ, σ)</i>		<i>N(μ, σ)</i>
<i>CDF.CHISQ(x, gl)</i>		<i>gl</i> askatasun-graduko khi karratu banaketa
<i>CDF.T(x, gl)</i>		<i>gl</i> askatasun-graduko Student-en banaketa
<i>CDF.F(x, gl₁, gl₂)</i>		<i>gl₁</i> eta <i>gl₂</i> askatasun- gradutako <i>F</i> banaketa
<i>IDF.UNIFORM(p, a, b)</i>	<i>P(X ≤ x) = p</i> probabilitatearen <i>x</i> balioa	<i>Uniform(a, b)</i>
<i>IDF.EXP(p, 1/β)</i>		<i>Exp(β)</i>
<i>IDF.NORMAL(p, μ, σ)</i>		<i>N(μ, σ)</i>
<i>IDF.CHISQ(p, gl)</i>		<i>gl</i> askatasun-graduko khi karratu banaketa
<i>IDF.T(p, gl)</i>		<i>gl</i> askatasun-graduko Student-en banaketa
<i>IDF.F(p, gl₁, gl₂)</i>		<i>gl₁</i> eta <i>gl₂</i> askatasun- gradutako <i>F</i> banaketa

4.2. Adibide batzuk

4.2.1. adibidea

Demagun lantegi bateko jardueren iraupenak banaketa uniformea duela, 12 orduko batez bestekoa eta 3 ordu karratuko bariantza edukiz.

- a) Kalkula bedi edozein jarduerak 12 ordu baino gutxiago irauteko probabilitatea.
- b) Lor bedi c konstantearen balioa, edozein jarduerak c ordu baino gehiago irauteko probabilitatea 0,85ekoa dela jakinik.

Ebazpena:

Biz $X = \text{"lantegiko jardueren iraupena"}$ aldagai aleatorio jarraitua, banaketa uniformekoa $[a, b]$ tartean.

$X: \text{Uniform}(a, b)$ aldagaiaren batez bestekoa eta bariantzaren balioak ezagunak direnez,

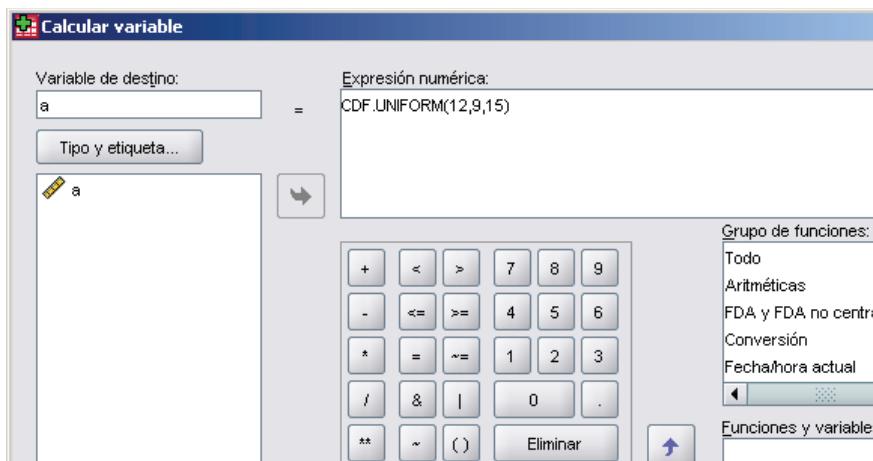
$$E(X) = \frac{a+b}{2} = 12 \quad Var(X) = \frac{(b-a)^2}{12} = 3$$

ekuazio-sistema ebatziz banaketaren a eta b parametroen balioak lortuko dira.

$$a = 9 \quad b = 15$$

Beraz, $X: \text{Uniform}(9, 15)$ da.

- a) 4.1. irudiak $P(X < 12)$ balioa nola kalkulatu adierazten du:



4.1. irudia.

$$P(X < 12) = CDF.UNIFORM(12, 9, 15) = 0,5$$

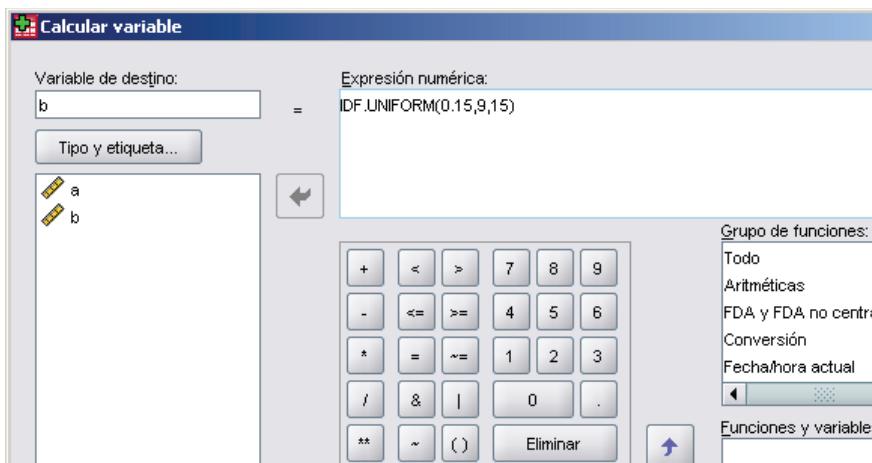
Erabilitako sintaxiaren emaitza 4.2. irudian jaso da:

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
	Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda							
1: a	,5000							
	a	b	var	var	var	var	var	var
1	5000	.						
2								
3								
4								

4.2. irudia.

b) *IDF* funtzioaren bidez lortuko da eskatutako balioa (ikus 4.3. irudia):

$$0,85 = P(X > c) \rightarrow 0,15 = P(X \leq c) = IDF. UNIFORM(0.15, 9, 15)$$



4.3. irudia.

Emaitzia 4.4 irudian jaso da, $c = 9,9$ delarik:

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
	Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda							
1 : b	9.90							
1	.5000	b	var	var	var	var	var	var
2								
3								
.								

4.4. irudia.

4.2.2. adibidea

Bira $X = \text{"LH motorraren hileko kontsumoa"}$ aldagai aleatorioa eta $f(x)$ dagokion dentsitate-funtzioa:

$$f(x) = \begin{cases} \frac{e^{-x/100}}{100} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

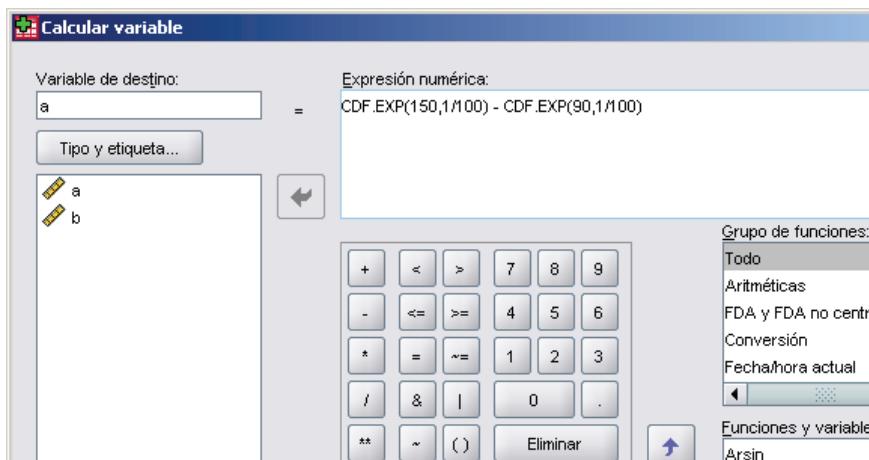
- a) Kalkulatu LH motorraren hileko kontsumoa 90 eta 150 litro bitartekoa izateko probabilitatea.

- b) Zein da motorraren hileko kontsumoa 80 litro baino handiagoa izateko probabilitatea?
- c) Lortu c konstantearen balioa, jakinda LH motorraren kontsumoa c baino txikiagoa dela 0,95eko probabilitatez.

Ebazpena:

X aldagaiari dagokion dentsitate-funtzioaren arabera, 100 parametroko banaketa esponentziala du X aldagai aleatorioak.

- a) Lehenbizi, atal honetan $P(90 \leq X \leq 150)$ kalkulatu behar da (4.5. irudia).



4.5. irudia.

4.6 irudian jaso da emaitza:

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda								
1: a .183439499592169								
1	a	b	var	var	var	var	var	var
1	.1834							
2								
3								
4								

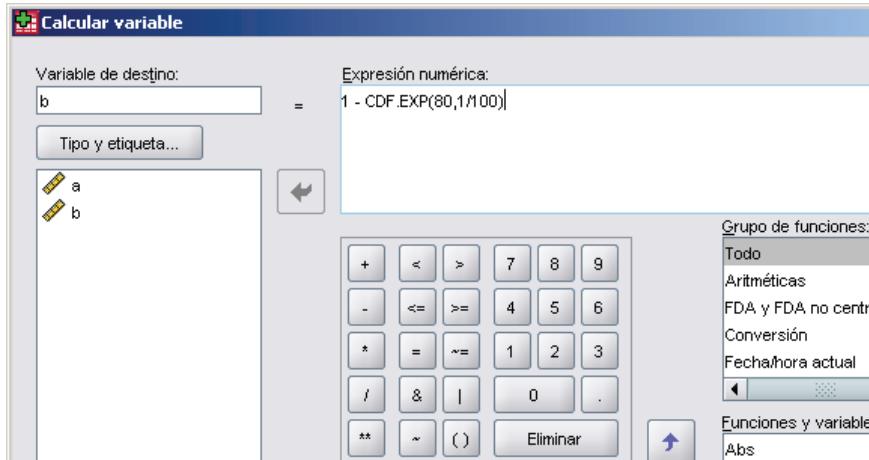
4.6. irudia.

Hau da:

$$P(90 \leq X \leq 150) = CDF.EXP(150, 1/100) - CDF.EXP(90, 1/100) = 0,1834$$

- b) $P(X > 80)$ balioa kalkulatzeko honela eragin daiteke (4.7. eta 4.8. irudiak):

$$P(X > 80) = 1 - CDF.EXP(80, 1/100) = 0,4493$$



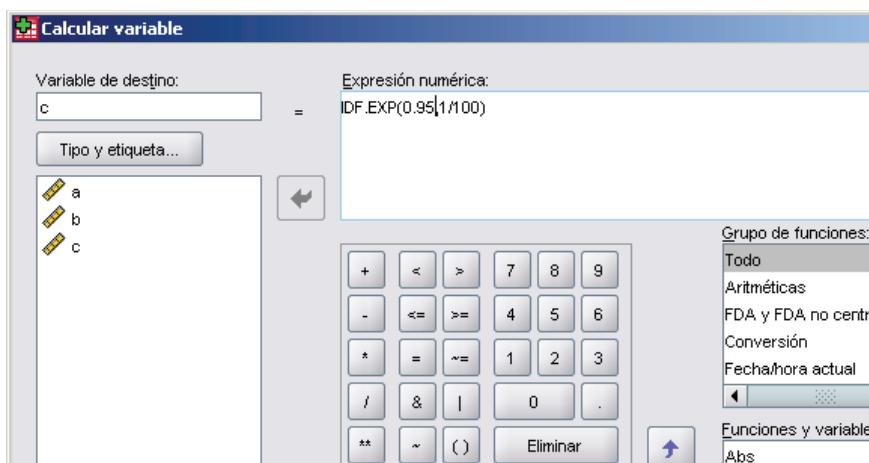
4.7. irudia.

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda								
1 : b .449328964117222								
	a	b	var	var	var	var	var	var
1	.1834	.4493						
2								
3								
4								

4.8. irudia.

- c) Oraingoan $P(X < c) = 0,95$ balioa ezaguna izanik, c konstantearen balioa lortzeko ondoko sintaxia erabil daiteke (4.9. irudia):

$$IDF.EXP(0.95, 1/100) = c$$



4.9. irudia.

Emaitza 4.10. irudian adierazi da:

$$c = 299,57$$

	a	b	c	var	var	var
1	.1834	.4493	299.57			
2						
3						
.						

4.10. irudia.

4.2.3. adibidea

Aldagai aleatorio batek $\mu = 600$ batez bestekodun banaketa normala du.

- Aldagaiak 500 baino txikiagoak diren balioak hartzeko probabilitatea 0,23 dela jakinda, kalkula bedi aldagai horren desbideratze tipikoa.
- Zein x baliok uzten du banaketaren %85 bere eskuinean?

Ebazpena:

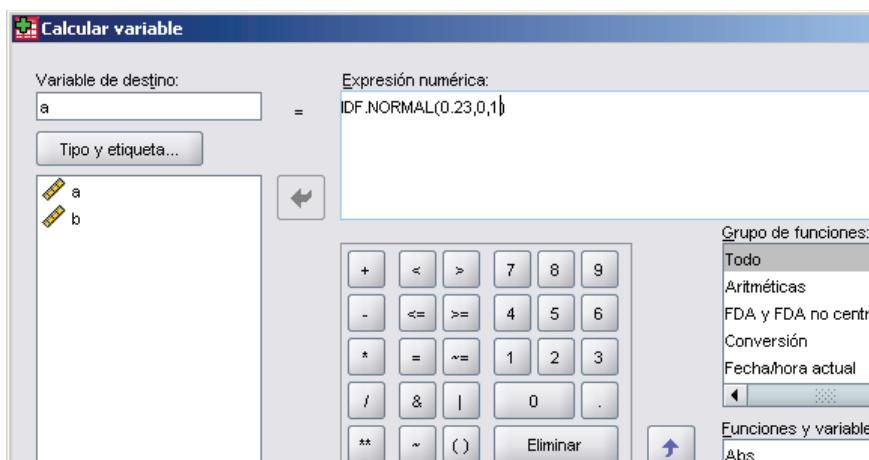
- a) Hasteko, enuntziatuko datuak hartu eta tipifikatu egingo da:

$$0,23 = P(X < 500) = P\left(\frac{X - \mu}{\sigma} < \frac{500 - 600}{\sigma}\right)$$

$Z = \frac{X - \mu}{\sigma}$ aldagai aleatorioak $N(0, 1)$ banaketa duenez,

$$IDF.NORMAL(0.23, 0, 1) = \frac{500 - 600}{\sigma}$$

ebatziz lortuko da aldagaiaren desbideratze tipikoa (4.11. irudia).



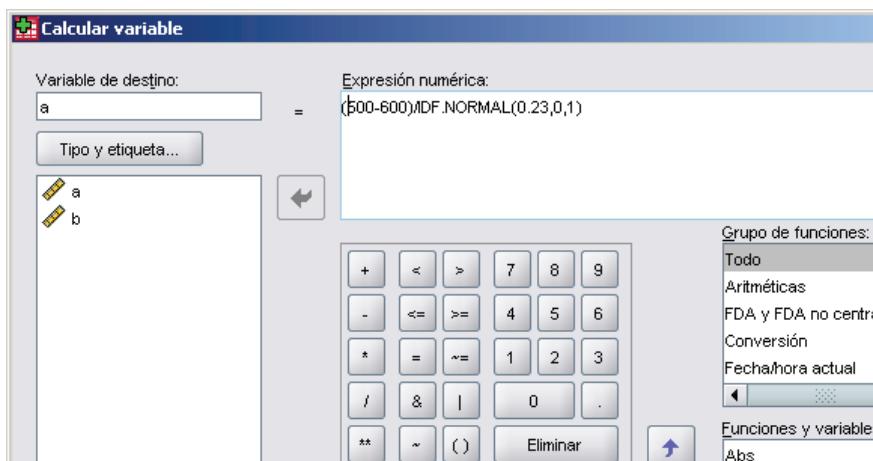
4.11. irudia.

4.12. irudian $\frac{500 - 600}{\sigma} = -0,74$ balioa jaso da.

The screenshot shows the SPSS Data Editor window titled '*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos'. The menu bar includes Archivo, Edición, Ver, Datos, Transformar, Analizar, Gráficos, Utilidades, Ventana, and Ayuda. The toolbar has various icons for file operations. The data view shows a single row labeled '1: a' with the value '-0.74' in the first column. The second column is labeled 'b' and contains the value '135.35'. The remaining columns are labeled 'var' and have empty cells.

4.12. irudia.

Desbideratze tipikoaren balioa lortzeko 4.13. irudian adierazi den prozedura erabiliko da eta emaitza 4.14. irudian ikus daiteke:



4.13. irudia.

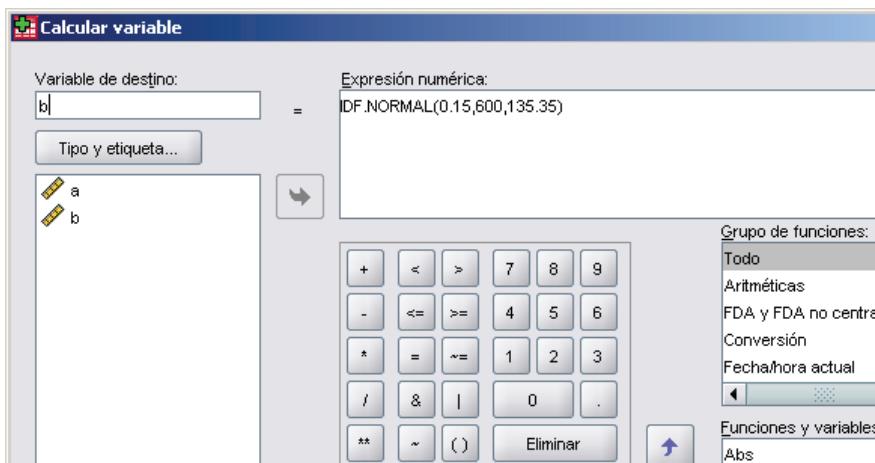
The screenshot shows the SPSS Data Editor window with the same data structure as before. The 'a' column now contains the value '135.35' in the first row, indicating the calculated mean.

4.14. irudia.

Beraz, $\sigma = 135,35$ da.

- b) Atal honetako x balioa lortzeko erabiliko den sintaxia (4.15. irudia) hauxe da:

$$IDF.NORMAL(0.15, 600, 135.35) = x$$



4.15. irudia.

4.16. irudian jaso da $x = 459,72$ emaitza:

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda								
1: b		459.7187407320156						
	a	b	var	var	var	var	var	v
1	135.35	459.72						
2								
3								
4								

4.16. irudia.

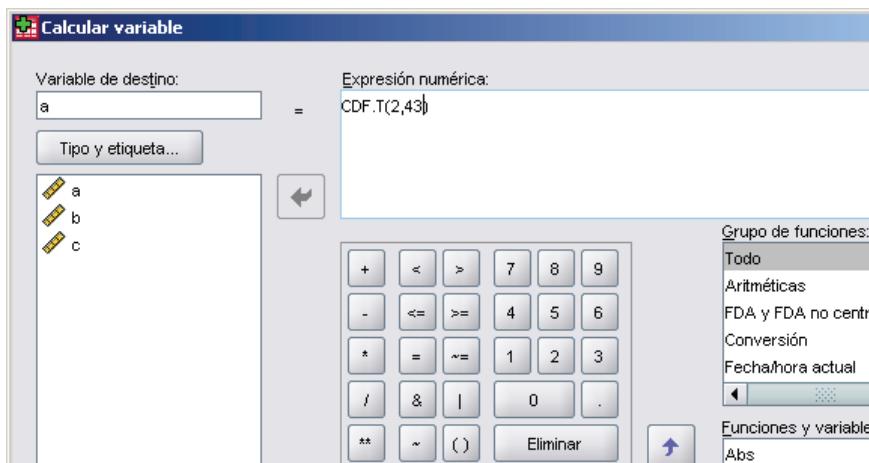
4.2.4 adibidea

Biz X aldagai aleatorioa, 43 askatasun-graduko Student-en banaketakoa. Lortu hurrengo balioak:

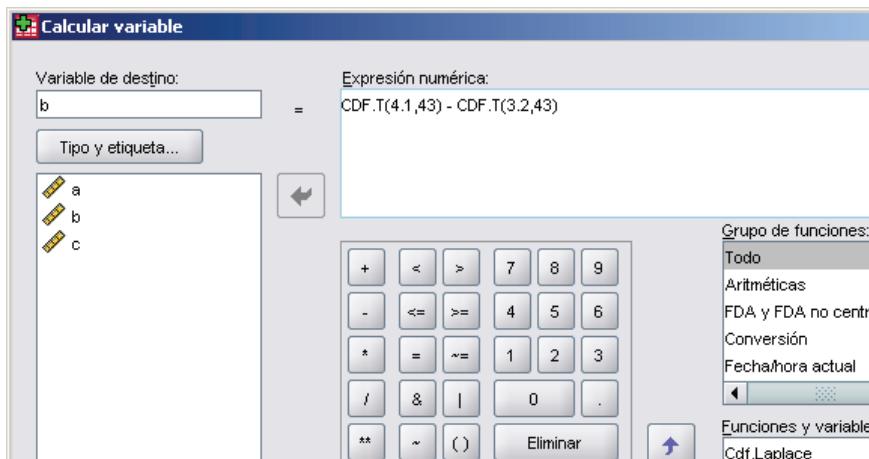
- a) $P(X < 2)$
- b) $P(3,2 < X < 4,1)$
- c) a balioa, $P(X > a) = 0,90$ dela jakinik.

Ebazpena:

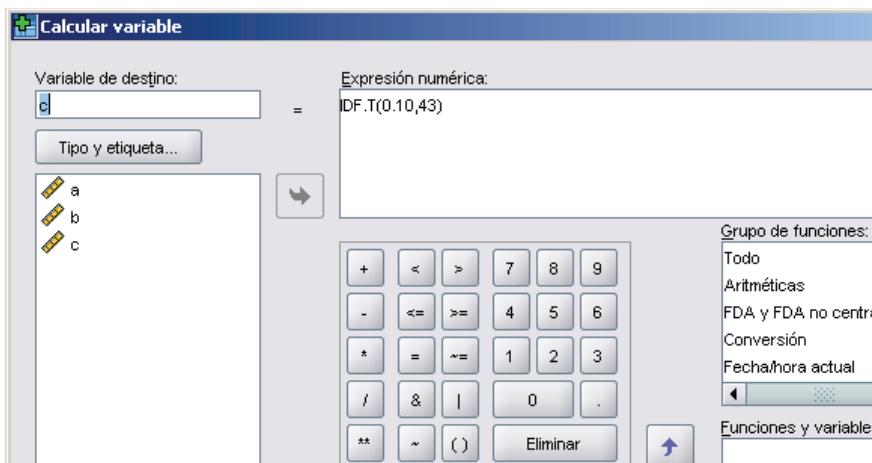
Abidide honetako atalak ebazteko erabili den sintaxia, 4.17. - 4.19. irudietan dago:



4.17. irudia.



4.18. irudia.



4.19. irudia.

Bukatzeko, 4.20. irudian adierazi dira emaitzak:

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos								
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda								
1: a 974078983151772								
	a	b	c	var	var	var	var	v
1	.9741	.0012	-1.30					
2								
3								
4								

4.20. irudia.

4.2.5 adibidea

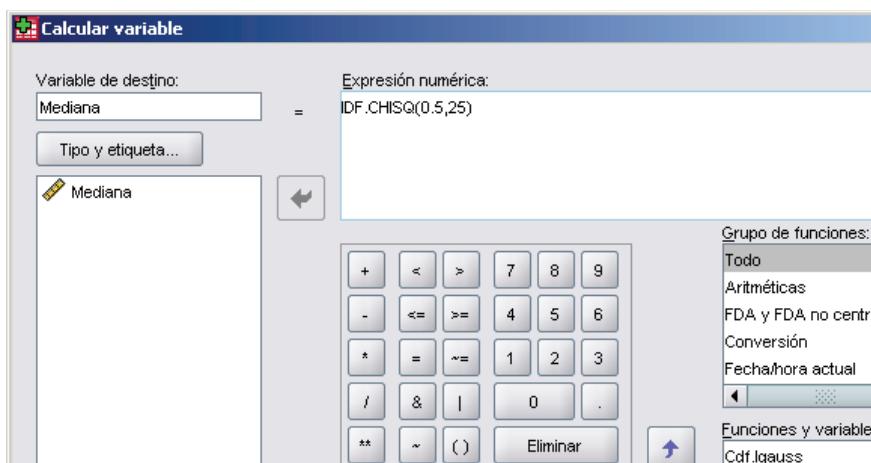
Kalkulatu 25 askatasun-graduko khi karratu banaketa duen aldagaiaren mediana.

Ebazpena:

Medianari dagokion banaketa-funtzioaren balioa 0,5 da. Beraz,

$$IDF.CHISQ(0.5,25)$$

funtzioaren bidez (4.21. irudia) lortuko da mediana:



4.21. irudia.

Kasu honetan medianaren balioa 24,34 da (ikus 4.22. irudia).

*AA.sav [Conjunto_de_datos1] - PASW Statistics Editor de datos	
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda	
1: Mediana 24.336586697884297	
1	Mediana 24.34
2	
3	
4	

4.22. irudia.

4.2.6. adibidea

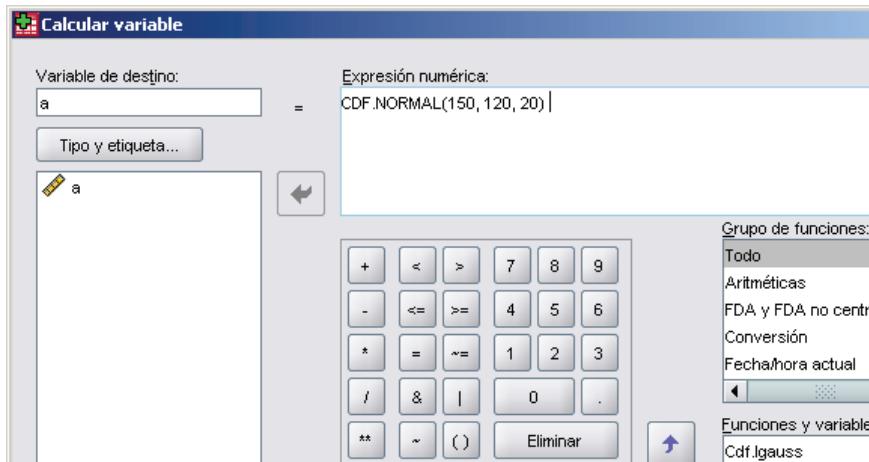
Makina-mota bat muntatzeko behar den denbora 2 orduko batez besteko eta 20 minutuko desbideratze tipikodun banaketa normala duen aldagai aleatorioa da. Zein da makina muntatzeko gehienez bi ordu eta erdi behar izateko probabilitatea?

Ebazpena:

Biz X = “makina-mota bat muntatzeko denbora” aldagai aleatorioa, $\mu = 120$ minuti eta $\sigma = 20$ minutuko banaketa normalekoa.

Problema ebazteko sintaxia eta emaitza hauexek dira:

$$P(X \leq 150) = CDF.NORMAL(150, 120, 20) = 0,0065$$



4.23. irudia

	a	var	var	var	var	var	var	v
1	.9332							
2								
3								
4								

4.24. irudia.

4.23. irudian sintaxia adierazi da eta 4.24. irudian emaitza jaso da.

4.3. Praktikatzeko ariketak

4.3.1. ariketa

Biz X = “bi konposatu kimikoren artean erreakzioa gertatzeko behar den denbora”. Demagun X aldagai aleatorioak 20 minutu eta 40 minutu bitarteko banaketa uniformea duela. Kalkula bedi erreakzioa gertatzeko 25 eta 35 minutu bitarteko

denbora behar izateko probabilitatea. Zein da erreakzioa gertatzeko gehienez ordu erdi behar izateko probabilitatea?

4.3.2. ariketa

Makina batean pieza bat muntatzeko behar den denbora $\mu = 25,5$ minutu eta $\sigma^2 = 5$ (minutu)² parametroetako banaketa normala duen aldagai aleatorioa da.

- Zein da piezaren muntaketa bukatzeko gutxienez hogei minutu behar izateko probabilitatea?
- Kalkula bedi piezaren muntaketa bukatzeko 22 minutu baino gehiago eta 28 minutu baino gutxiago behar izateko probabilitatea.

4.3.3. ariketa

Demagun osagai mota berezi baten iraupena urtetan aztertu nahi dela. Jakinda aldagaiari dagokion dentsitate-funtzioa

$$f(x) = \begin{cases} \frac{e^{-x/4}}{4}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

dela, kalkulatu osagaietako bi urte eta lau urte bitarteko iraupena izateko probabilitatea. Zein da osagaietako gutxienez urte bateko iraupena izateko probabilitatea?

4.3.4. ariketa

Kalkulatu hurrengo balioak:

- a) $t_{0,05;11}$ b) $F_{0,95;12,10}$ c) $\chi^2_{0,025;30}$ d) $P(t_6 \geq 3,062)$ e) $P(\chi^2_4 < 9,49)$

5. ESTIMAZIOA

5.1. Sarrera

Gai honetan zenbait kasu praktikotako konfiantza-tartea zehaztuko dira, SPSS erabiliz. Lehenik, estimatu nahi den populazioko parametrorako konfiantza-tartea zein den jakin beharko da. Gero, tartaren muturren adierazpenak finkatuta, SPSS softwareean *Transformar > Calcular variable* aukeratuko da. Ondoren, *Variable de destino* koadroan aldagai berriaren izena idatzi edo hautatuko da, eta bukatzeko, *Expresión numérica* koadroan dagokion sintaxia idatziko da. *Aceptar* botoian klik egin ondoren, *Ventana de datos* delako leihoa SPSS programak lortu nahi diren konfiantza-tarteko bi muturrak adieraziko ditu.

5.2. Adibide batzuk

5.2.1. adibidea

Demagun inprimagailuko tonerraren iraupena neurten duen aldagai aleatorioak banaketa normala duela. Aleatorioki 20 toner hartu dira eta hauen batez besteko iraupena 92 egun eta desbideratze tipikoa 7 egun izan dira. Estima bedi tonerraren batez besteko iraupena %95eko konfiantza-mailaz. Eta %99ko konfiantza-mailaz?

Ebazpena

Biz X = “inprimagailuko tonerraren iraupena” aldagai aleatorioa μ eta σ parametroetako banaketa normalekoa. Lagin aleatorio bakuneko datuak $n = 20$ toner, $\bar{x} = 92$ egun eta $s = 7$ egun dira.

- a) Bariantza ezezaguneko populazioaren batez bestekoa estimatzeko, laginaren tamaina $n < 30$ denean, hurrengo konfiantza-tartea eraikiko da:

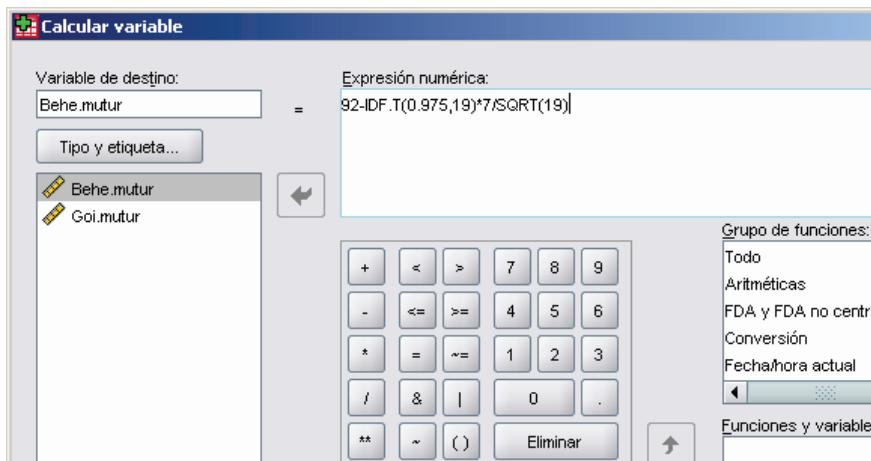
$$I_{\mu}^{1-\alpha} = \left[\bar{x} - t_{\alpha/2;n-1} \frac{s}{\sqrt{n-1}}, \bar{x} + t_{\alpha/2;n-1} \frac{s}{\sqrt{n-1}} \right]$$

Ondoren, enuntziatuko datuak ordezkatuko dira, $1 - \alpha = 0,95$ konfiantza-maila hartuz:

$$I_{\mu}^{0,95} = \left[92 - t_{0,025;19} \frac{7}{\sqrt{19}}, 92 + t_{0,025;19} \frac{7}{\sqrt{19}} \right]$$

Tarte horretako behe-muturra kalkulatzeko, SPSS programan *Transformar > Calcular variable* aukeratuko da. 5.1. irudian ikus daitekeen modura, *Variable de destino* koadroan *Behe.mutur* izeneko aldagai hautatu eta hurrengo sintaxia idatziko da:

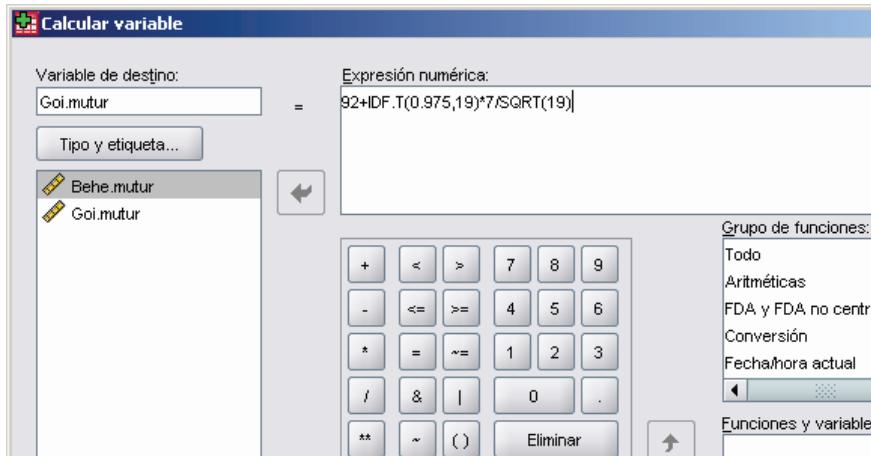
$$92 - IDF.T(0.975,19) * 7 / SQRT(19)$$



5.1. irudia.

Urrats berdinak emanez lor daiteke goi-muturra (5.2. irudia), honako sintaxia erabiliz:

$$92 + IDF.T(0.975,19) * 7 / SQRT(19)$$



5.2. irudia.

SPSS programako datuen leihoa daude konfiantza-tarteko bi muturrak, 5.3. irudian adierazten den modura:

	Behe.mutur	Goi.mutur	var	var	var	var	var	v
1	88.64	95.36						
2								
3								
4								

5.3. irudia.

Hau da, konfiantza-tartea $I_{\mu}^{0,95} = [88,64, 95,36]$ da. Ondorioz, %95eko konfiantza-mailaz, inprimagailuko tonerraren batez besteko iraupena 89 eta 95 egunekoa bitartekoa dela baiezta daiteke.

Konfiantza-maila %99 bada, kalkuluak errepikatu beharko dira $1 - \alpha = 0,99$ hartuz. Kasu horretan honako konfiantza-tartea zehaztuko da:

$$I_{\mu}^{0,99} = \left[92 - t_{0,005;19} \frac{7}{\sqrt{19}}, 92 + t_{0,005;19} \frac{7}{\sqrt{19}} \right]$$

Behe-muturra eta goi-muturraren balioak lortzeko, hurrengo sintaxia erabiliko da, hurrenez hurren:

$$92 - IDF.T(0.995,19) * 7 / SQRT(19)$$

$$92 + IDF.T(0.995,19) * 7 / SQRT(19)$$

Konfiantza-tartea hauxe da:

$$I_{\mu}^{0,99} = [87,41, 96,59]$$

Ondorioz, %99ko konfiantza-mailaz, inprimagailuko tonerraren batez besteko iraupena gutxienez 87 egunekoa eta gehienez 97 eguneko dela onar daiteke.

5.2.2. adibidea

Elektronika alorrean lan egiten duten emakumeen proportzioa estimatzeko, aleatorioki elektronika sailetako 180 langile hartu dira eta hauetatik 90 emakumeak dira. %95eko konfiantza-mailaz, zehaztu zein tartetan koka daitekeen elektronika alorrean lan egiten duen emakumeen proportzioa eta kalkulatu estimazio-errorea.

Ebazpena

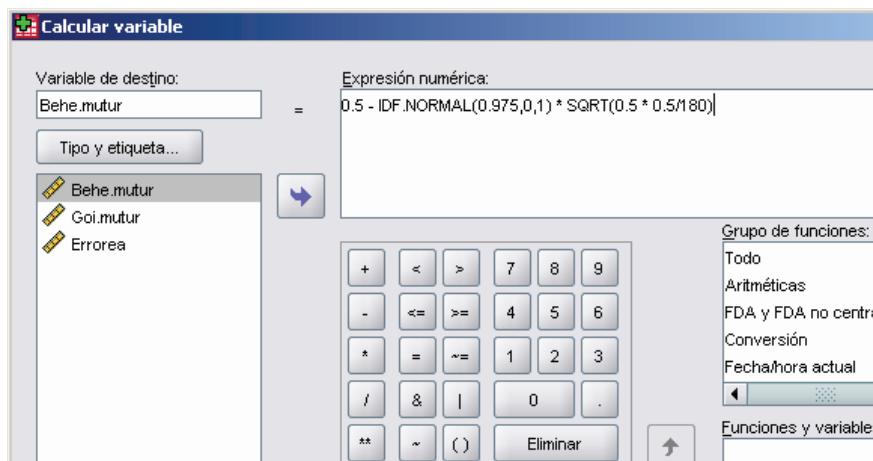
Bira p = “elektronika alorreko emakumeen proportzioa”, $n = 180$ laginaren tamaina eta $\hat{p} = \frac{90}{180} = 0,5$ emakumeen proportzioa laginean. Kasu honetan, populazioaren batez bestekoa estimatzeko konfiantza-tartearen adierazpen orokorra hurrengoa da:

$$I_p^{0,95} = \left[0,5 - z_{0,025} \sqrt{\frac{0,5 \cdot 0,5}{180}}, 0,5 + z_{0,025} \sqrt{\frac{0,5 \cdot 0,5}{180}} \right]$$

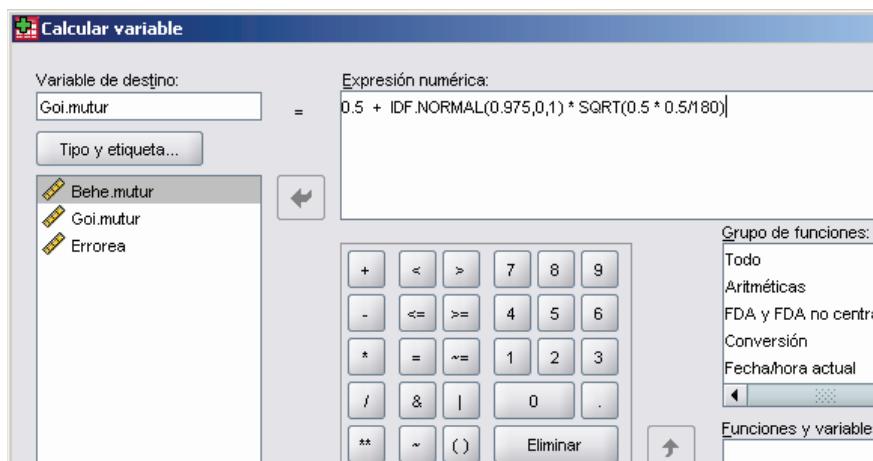
non $1 - \alpha = 0,95$ den.

Konfiantza-tarte horretako bi muturrak zehazteko, 5.4. eta 5.5. irudietan adierazi moduan, *Transformar > Calcular variable* prozeduraren barruan SPSSren bidezko hurrengo sintaxia erabiliko da:

$$\begin{aligned} 0.5 - IDF.NORMAL(0.975, 0,1) * SQRT(0.5 * 0.5 / 180) \\ 0.5 + IDF.NORMAL(0.975, 0,1) * SQRT(0.5 * 0.5 / 180) \end{aligned}$$



5.4. irudia.



5.5. irudia.

Horrela, $I_p^{0.95} = [0,427, \quad 0,573]$ konfiantza-tartea lortuko da (5.6. irudia).

	Behe.mutur	Goi.mutur	Errorea	var	var	var	var	V
1	.427	573	.					
2								
3								

5.6. irudia.

Beraz, elektronika sailetako emakumeen proportzioa %42,7 eta %57,3 bitartekoa da, %95eko konfiantza-mailaz.

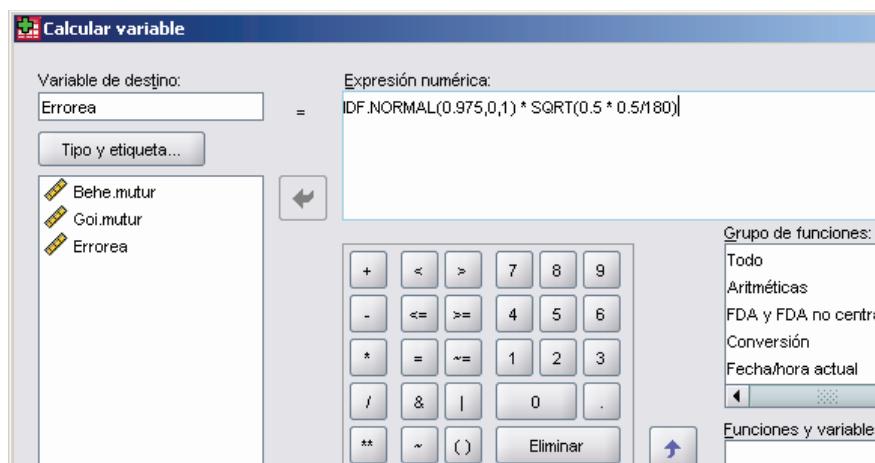
Estimazio-errorea zehazteko honako kalkulua egingo da:

$$\varepsilon = z_{0,025} \sqrt{\frac{0,5 \cdot 0,5}{180}}$$

SPSS programan *Transformar > Calcular variable* aukeratu ondoren,

$$IDF.NORMAL(0.975,0,1) * SQRT(0.5 * 0.5 / 180)$$

agindua emango da (5.7. irudia), eta estimazio-errorearen balioa ($\varepsilon = 0,073$) lortuko da (5.8. irudia):



5.7. irudia.

*Sin título1 [Conjunto_de_datos0] - PASW Statistics Editor de datos

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda

1 : Errorea .073043545048048

	Behe.mutur	Goi.mutur	Errorea	var	var	var	var
1	.427	.573	.073				
2							
3							
4							

5.8. irudia.

5.2.3. adibidea

Makina paketatzale batek azukre paketeak betetzen ditu. Demagun paketeko azukre-kantitatea neurten duen aldagai aleatorioak banaketa normala duela. Aleatorioki 30 pakete azukre hartu dira, zeintzuen azukre-kantitatearen desbideratze tipikoa gramo batekoa den. Lor bedi azufre-kantitatearen desbideratze tipikoa estimatzeko %99ko konfiantza-mailako konfiantza-tartea.

Ebazpena

Biz X = "azukre-kantitatea paketeko" banaketa normala duen aldagai aleatorioa. Lortu behar den konfiantza-tartea hauxe da:

$$I_{\sigma}^{1-\alpha} = \left[\frac{s\sqrt{n}}{\sqrt{\chi^2_{\alpha/2;n-1}}}, \frac{s\sqrt{n}}{\sqrt{\chi^2_{1-\alpha/2;n-1}}} \right]$$

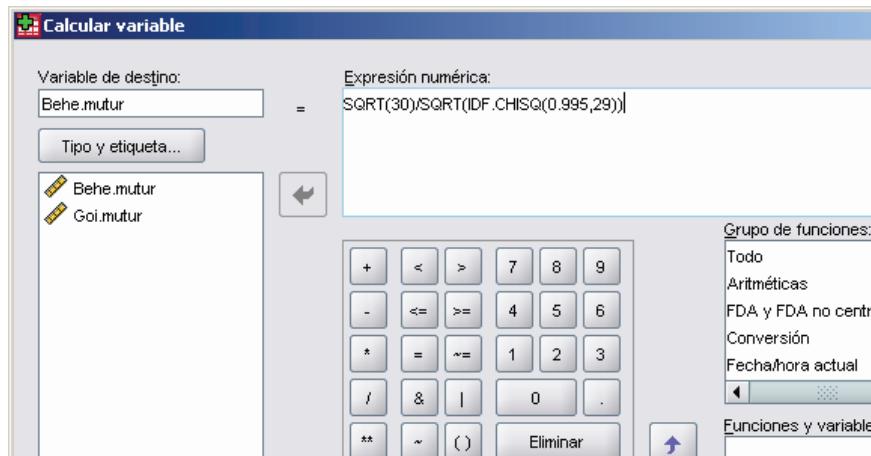
Konfiantza-tartean $n = 30$, $s = 1$ eta $1 - \alpha = 0,99$ balioak ordezkatz, hurrengoa lortuko da:

$$I_{\sigma}^{0,99} = \left[\frac{1 \cdot \sqrt{30}}{\sqrt{\chi^2_{0,005;29}}}, \frac{1 \cdot \sqrt{30}}{\sqrt{\chi^2_{0,995;29}}} \right]$$

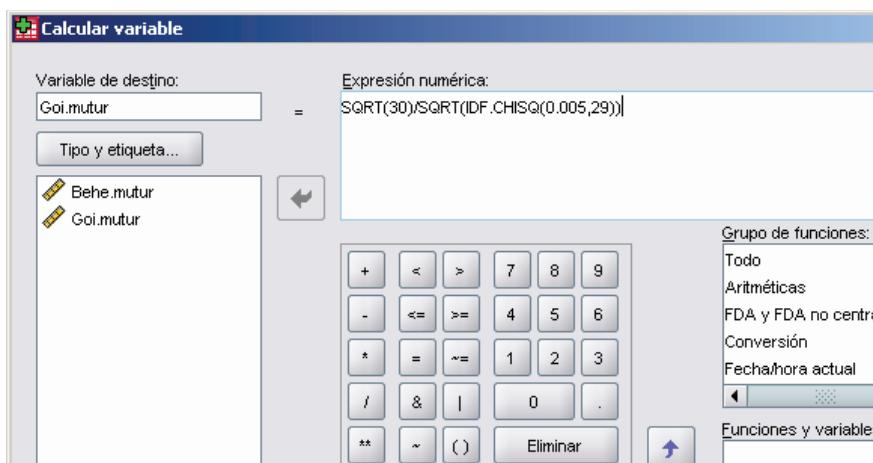
Behe-muturra eta goi-muturraren balioak zehazteko, *Transformar > Calcular variable* prozeduran (5.9. eta 5.10. irudiak) honako sintaxia erabiliko da, hurrenez hurren:

$SQRT(30) / SQRT(IDF.CHISQ(0.995,29))$

$SQRT(30) / SQRT(IDF.CHISQ(0.005,29))$



5.9. irudia.



5.10.irudia.

Ondorioz, hurrengo konfiantza-tartea (5.11. irudia)lortuko da:

$$I_{\sigma}^{0.99} = [0,76, 1,51]$$

* Sin título1 [Conjunto_de_datos0] - PASW Statistics Editor de datos								
	Archivo		Edición		Ver		Datos	
1: Behe.mutur	.75115173087113		var	var	var	var	var	V
1	Behe.mutur	.76	Goi.mutur	1.51				
2								
3								
4								

5.11.irudia.

5.2.4 adibidea

Lantegi batean ekoizten den A motako pieza-kantitatea eta B motako pieza-kantitatea neurten dituzten aldagai aleatorioak elkarrekiko independenteak eta banaketa normalekoak dira. Aleatorioki 31 egunetako ekoizpena hartu da eta A motako batez besteko pieza-kantitatea 100 eta desbideratze tipikoa 10 izan dira. Aleatorioki hartutako egun horietan B motako batez besteko pieza-kantitatea 90 eta desbideratze tipikoa 9 izan dira.

- a) %90eko konfiantza mailaz, lortu batez besteko ekoizpenen arteko diferentzia estimatzeko konfiantza-tartea.
- b) %90eko konfiantza mailaz, zein da bariantzen arteko zatidura estimatzeko konfiantza-tartea?

Ebazpena

Bira X_1 = “A motako pieza-kantitatea”, X_2 = “B motako pieza-kantitatea” aldagai aleatorioak.

$$n = 31 \text{ egun}, \quad \bar{x}_1 = 100 \text{ pieza}, \quad s_1 = 10 \text{ pieza}$$

$$n = 31 \text{ egun}, \quad \bar{x}_2 = 90 \text{ pieza}, \quad s_2 = 9 \text{ pieza}$$

- a) Kasu honetan, bariantza ezezagunetako eta banaketa normaleko bi aldagai aleatorio independente ditugunez eta laginetako tamainak 30 baino handiagoak direnez, batez besteko arteko diferentzia estimatzeko konfiantza-tartea hauxe da:

$$I_{\mu_1 - \mu_2}^{1-\alpha} = \left[(\bar{x}_1 - \bar{x}_2) \mp z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1-1} + \frac{s_2^2}{n_2-1}} \right]$$

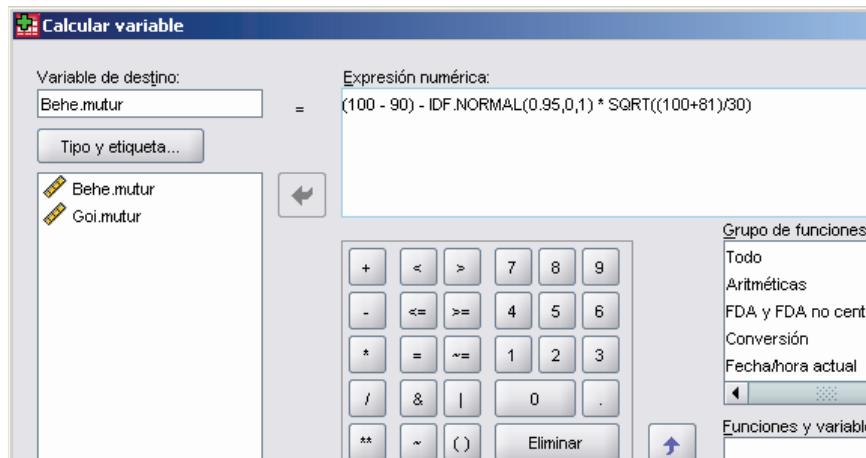
Konfiantza-tartean $1 - \alpha = 0,90$ ordezkatuko da.

$$I_{\mu_1 - \mu_2}^{0,90} = \left[(100 - 90) \mp z_{0,05} \sqrt{\frac{10^2}{31-1} + \frac{9^2}{31-1}} \right]$$

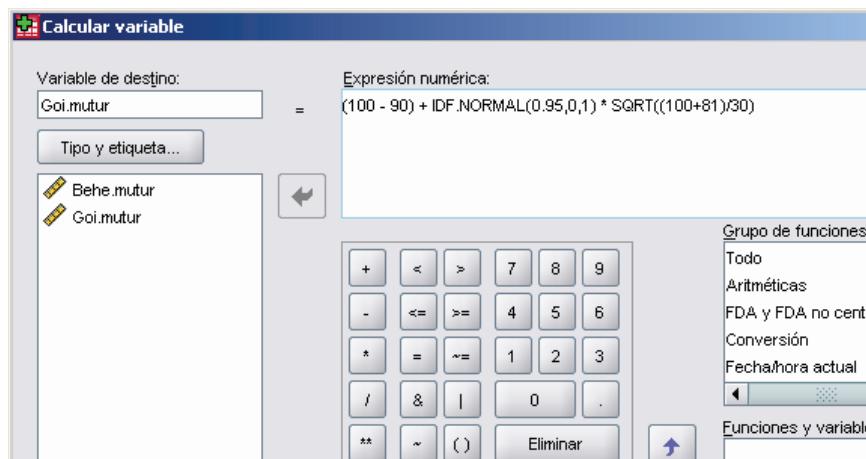
Behe-muturra eta goi-muturraren balioak zehazteko, *Transformar > Calcular variable* prozeduruan (5.12. eta 5.13. irudiak) hurrengo sintaxia erabiliko da, hurrenez hurren:

$$(100 - 90) - IDF.NORMAL(0.95,0,1) * SQRT((100 + 81)/30)$$

$$(100 - 90) + IDF.NORMAL(0.95,0,1) * SQRT((100 + 81)/30)$$



5.12. irudia.



5.13. irudia.

Lortuko den konfiantza-tartea hauxe da:

$$I_{\mu_1 - \mu_2}^{0,90} = [5,96, 14,04]$$

	Behe.mutur	Goi.mutur	var	var	var	var	var	v
1	5.96	14.04						
2								
3								

5.14. irudia.

b) Oraingoan lortu behar den konfiantza-tartea honakoa da:

$$I_{\sigma_1^2 / \sigma_2^2}^{1-\alpha} = \left[\frac{S_1^2 / S_2^2}{F_{\alpha/2; n_1-1, n_2-1}}, \frac{S_1^2 / S_2^2}{F_{1-\alpha/2; n_1-1, n_2-1}} \right],$$

non $1 - \alpha = 0,90$ ordezkatuko den.

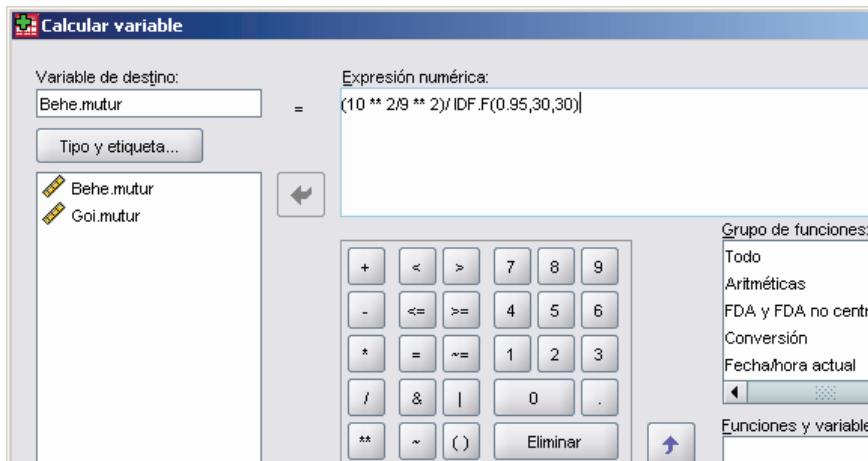
Gainera, kasu honetan laginak tamaina berekoak direnez, laginen kuasibariantzen arteko zatidura bariantzen arteko zatiduraren berdina da.

$$\frac{S_1^2}{S_2^2} = \frac{\frac{n}{n-1} s_1^2}{\frac{n}{n-1} s_2^2} = \frac{s_1^2}{s_2^2}$$

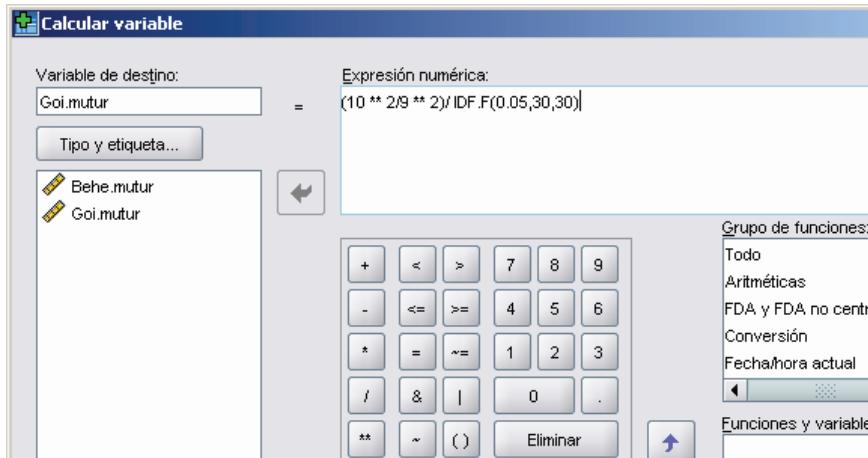
Bukatzeko, behe-muturra eta goi-muturraren balioak zehazteko *Transformar > Calcular variable* prozeduran (5.15. eta 5.16. irudiak), hurrengo sintaxia erabiliko da, hurrenez hurren:

$$(10^{**2}/9^{**2})/IDF.F(0.95,30,30)$$

$$(10^{**2}/9^{**2})/IDF.F(0.05,30,30)$$



5.15. irudia.



5.16. irudia.

Ondorioz, eskatutako kofiantza-tartea (5.17. irudia) lortuko da:

$$I_{\sigma_1^2/\sigma_2^2}^{0.90} = [0.67, 2.27]$$

	Behe.mutur	Goi.mutur	var	var	var	var	var	V
1	67	2.27						
2								
3								

5.17. irudia.

5.2.5 adibidea

Makina batek goizeko txandan eta arratsaldeko txandan ekoizten dituen pieza akastunen proportzioak aztertu nahi dira. Aleatorioki goizeko txandan ekoitzitako 300 pieza hartu dira, hauetatik 18 pieza akastunak direlarik. Era berean, aleatorioki arratsaldeko txandan ekoitzitako 280 pieza aukeratu dira, 14 pieza akastunak direlarik. Demagun goizeko txandan ekoizten den pieza akastunen kopurua arratsaldeko txandan ekoizten denarekiko independentea dela. %95eko konfiantza-mailaz, lor bedi proportzioen differentziarako konfiantza-tartea.

Ebazpena

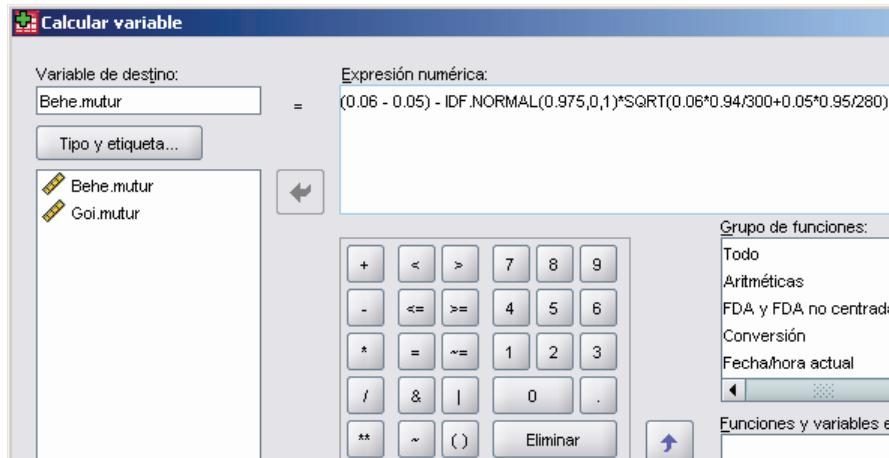
Bira p_1 = “goizeko pieza akastunen proportzioa”, p_2 = “arratsaldeko pieza akastunen proportzioa” eta $\hat{p}_1 = \frac{18}{300} = 0,06$ eta $\hat{p}_2 = \frac{14}{280} = 0,05$ hartutako lagin aleatorio bakunetako pieza akastunen proportzioak. Bi proportzioen arteko differentzia estimatzeko $1 - \alpha = 0,95$ konfiantza-mailako konfiantza-tartea hurrengoa da:

$$I_{p_1-p_2}^{1-\alpha} = \left[(\hat{p}_1 - \hat{p}_2) \mp z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \right]$$

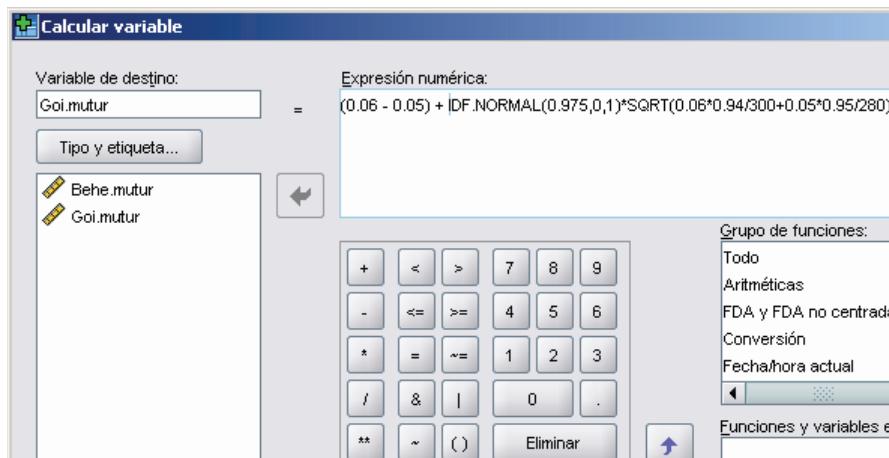
Konfiantza-tarteko muturrak kalkulatzeko, *Transformar > Calcular variable* prozeduran (5.18. eta 5.19. irudiak) sintaxi hauxe erabiliko da:

$$(0.06 - 0.05) - IDF.NORMAL(0.975,0,1) * SQRT\left(\frac{0.06 * 0.94}{300} + \frac{0.05 * 0.95}{280}\right)$$

$$(0.06 - 0.05) + IDF.NORMAL(0.975,0,1) * SQRT\left(\frac{0.06 * 0.94}{300} + \frac{0.05 * 0.95}{280}\right)$$



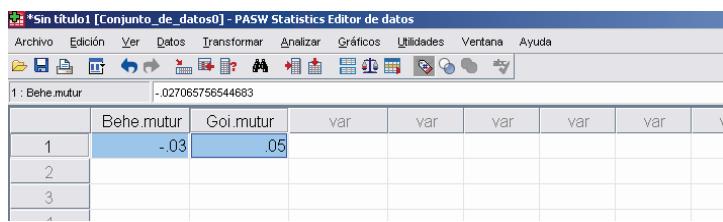
5.18. irudia.



5.19. irudia.

Ondorioz, hurrengo konfiantza-tartea (5.20. irudia) lortuko da:

$$I_{p_1-p_2}^{0,95} = [-0,03, 0,05]$$



	Behe.mutur	Goi.mutur	var	var	var	var	var	Y
1		-0.03	.05					
2								
3								
4								

5.20. irudia.

Beraz, %95eko konfiantza-mailaz proportzioen arteko diferentzia nula dela onar daiteke.

5.2.6 adibidea

Demagun fabrikazio-prozesu baten iraupena neurten duen aldagai aleatorioak banaketa normala duela. Aleatorioki n fabrikazio-prozesu hartu dira eta hauen batez besteko iraupena 8 ordu eta desbideratze tipikoa 2 ordu izan dira. %99ko konfiantza-mailaz, zehaztu zenbat fabrikazio-prozesu hartu behar diren orokorreko iraupena eta lagineko iraupenaren arteko diferentzia gehienez ordu betekoa izateko.

Ebazpena

Biz X = “fabrikazio-prozesuaren iraupena” aldagai aleatorioa $N(\mu, \sigma)$ banaketakoa. Alde batetik, populazioaren batez bestekoa estimatzeko %99ko konfiantza-mailako konfiantza-tartetik hurrengoa ondoriozta daiteke:

$$0,99 = P\left(\bar{x} - z_{0,005} \cdot s / \sqrt{n-1} \leq \mu \leq \bar{x} + z_{0,005} \cdot s / \sqrt{n-1}\right)$$

$s = 2$ delarik.

Hau da:

$$0,99 = P(|\bar{x} - \mu| \leq z_{0,005} \cdot 2 / \sqrt{n-1})$$

Bestalde, enuntziatuko testuak dioen arabera:

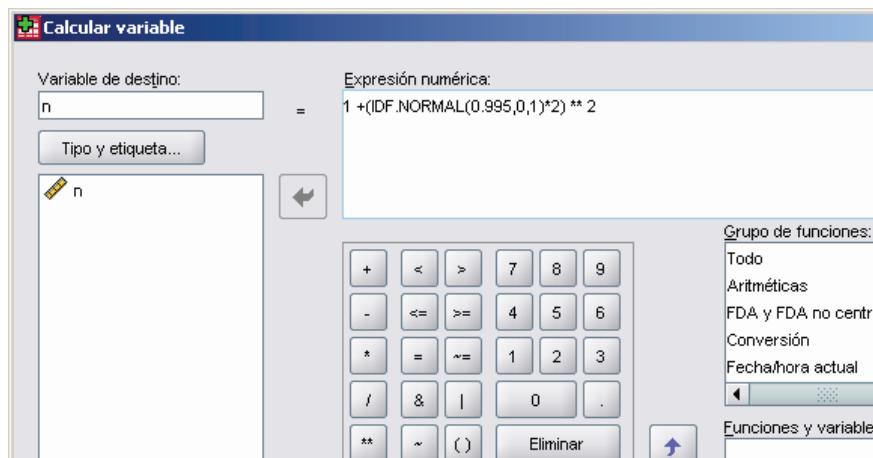
$$0,99 = P(|\bar{x} - \mu| \leq 1)$$

Beraz, %99ko konfiantza-mailaz hurrengoa ondoriozta daiteke:

$$1 = z_{0,005} \cdot 2 / \sqrt{n-1}$$

Azken ekuazio honetatik n bakandu behar da. Horrela, SPSS programan *Transformar > Calcular variable* prozeduraren koadroan (5.21. irudia) hurrengo sintaxia erabiliz, laginaren tamaina (5.22. irudia) lortuko da:

$$n = 1 + (\text{IDF.NORMAL}(0.995, 0, 1)^{*} 2)^{**} 2 \approx 28$$



5.21. irudia.

	n	var						
1	28							
2								
3								
.								

5.22. irudia.

5.3. Praktikatzeko ariketak

5.3.1. ariketa

Demagun pila alkalinoen iraupena banaketa normalekoa dela. Aleatorioki 18 pila hartu dira, batez besteko iraupena 10.500 ordu eta kuasibariantza 2.304 (ordu)² delarik.

- a) %97ko konfiantza-mailaz, zehatz bedi pila alkalinoen batez besteko iraupena.
- b) %97ko konfiantza-mailaz, estima bedi pila alkalinoen iraupenaren bariantza.

5.3.2. ariketa

Bi herrialdetako energia elektrikoaren kontsumoa aztergai da. Bi herrialdeetan urteko energia elektrikoaren kontsumoak banaketa normala du, A herrialdeko energia elektrikoaren kontsumoak 110 milioi kWh-ko desbideratze tipikoa eta B herrialdean 100 milioi kWh-ko desbideratze tipikoa izanik. Demagun energia elektrikoaren kontsumoak elkarrekiko independenteak direla. Azken 15 urteetan kontsumitutako energia elektrikoa neurtu da eta A herrialdean batez beste 2.356 milioi kWh eta B herrialdean batez beste 2.257 milioi Kwh kontsumitu dira. %99ko konfiantza-mailaz, onar al daiteke A herrialdeko energia elektrikoaren kontsumoa B herrialdekoa baino handiagoa izatea?

5.3.3. ariketa

Termometro batek orduko batez besteko tenperaturak neurtzen ditu, baina eguneko zenbait ordutan izandako tenperaturak falta dira (balio galduzat hartzen dira). Aleatorioki hartutako udako bi egunetan jasotako orduko batez besteko temperaturetatik 6 balio galdu dira. Era berean, aleatorioki hartutako neguko bi egunetan neurtutako orduko batez besteko temperaturetatik 8 balio galdu dira.

- a) %90eko konfiantza-mailaz, estima bedi udako temperaturen artean balio galduen proportzioa.
- b) %90eko konfiantza-mailaz, estima bedi udako balio galduen proportzioa eta neguko balio galduen proportzioaren arteko diferentzia.

6. HIPOTESI-KONTRASTEAK

6.1. Zenbait kontzeptu hipotesi-kontraste parametrikoan

Kasu praktiko askotan populazioaren parametro ezezagun bati buruzko hipotesiak onartu ala errefusatu erabaki behar da. Kasu horietan **hipotesi-kontraste parametrikoak** erabiltzen dira. Kontrastean **hipotesi nulua** (onartzeko ala errefusatzeko den hipotesia, orokorrean) eta **hipotesi alternatiboa** (beste hipotesia) finkatuko dira eta hartuko den erabakia α adierazgarritasun-maila jakin batekin egingo da. **Adierazgarritasun-maila** da: hipotesi nulua errefusatzeko probabilitatea, hipotesi nulua egia izanik. Hipotesi nulua onartzeko adierazgarritasun-maila maximoari **p-balioa** deritzo. Ondorioz, p-balioaren erregela aplikatuz hurrengo erabakiak har daitezke:

$$\begin{aligned}\alpha > p\text{-balioa} &\Rightarrow \text{errefusatu hipotesi nulua} \\ \alpha \leq p\text{-balioa} &\Rightarrow \text{onartu hipotesi nulua}\end{aligned}$$

Hurrengo adibideetan p-balioaren erregelan oinarrituz, SPSSren bidez har daitezkeen zenbait erabaki estatistiko azalduko dira. Horrela, lehenengo bi adibideetan SPSS programak eskaintzen dituen **T probak** azalduko dira. Proba hauek populazioaren batez bestekoari buruz edo bi populazioen batez bestekoen diferenciari buruzko hipotesi-kontrasteak egiteko erabil daitezke. Hirugarren adibidean normaltasunari buruzko kontraste ez-parametrikoa egingo da.

Bukatzeko, laugarren adibidean populazioaren proportzioari buruzko hipotesi-kontraste bat aztertuko da.

6.2. Adibide batzuk

6.2.1. adibidea

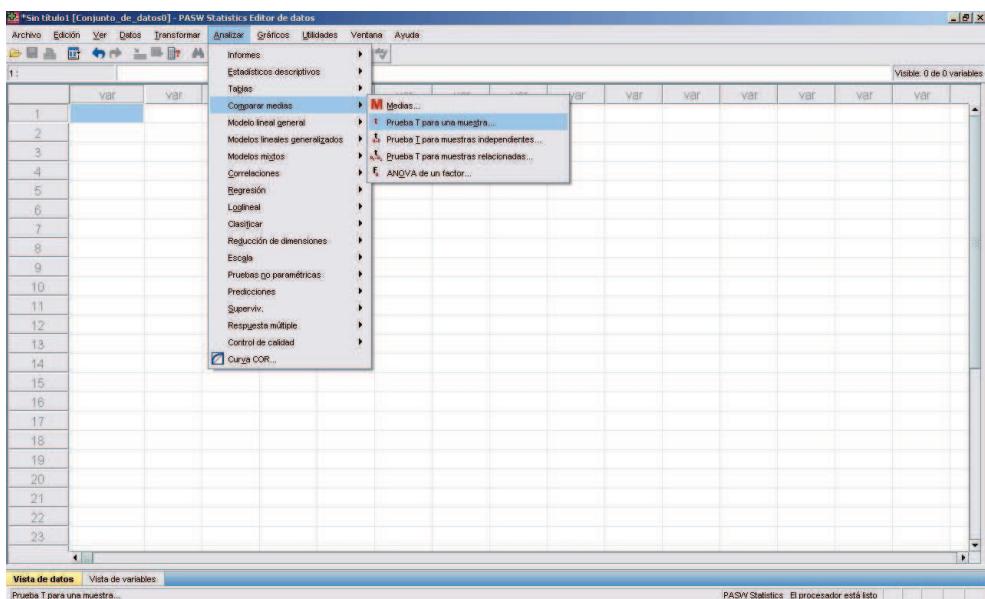
Demagun asteko ekoizpena neurten duen aldagai aleatorioak banaketa normala duela. Aleatorioki 60 astetan ekoitzi den produktu-kantitatea hurrengoa izan da:

101	116	103	98	97	118	118	104	107	106
97	111	89	119	105	120	127	112	106	117
92	100	98	114	107	91	109	98	121	109
99	121	106	129	94	104	129	114	107	108
97	116	108	99	101	130	111	106	111	105
104	131	95	85	136	101	113	111	101	106

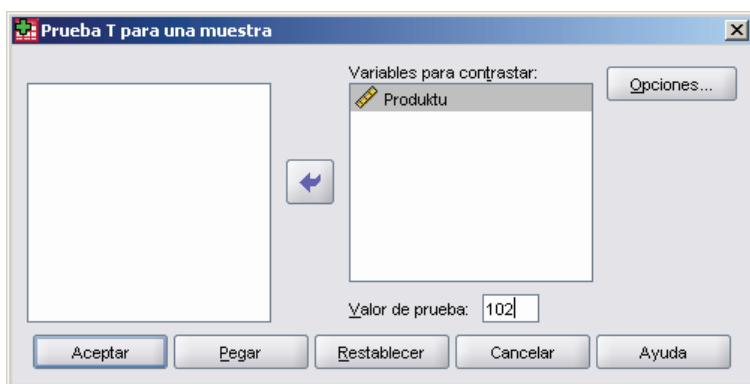
Erabili T proba, asteko batez besteko ekoizpena 102 produktukoa dela dioen hipotesia onartuko denetz erabakitzeko, %5eko adierazgarritasun-mailaz. Lortu asteko batez besteko ekoizpena estimatzeko %95eko konfiantza-mailako konfiantza-tartea.

Ebazpena:

Aldagaia banaketa normalekoa denez, T proba erabili ahal izango da. *Analizar > Comparar medias > Prueba T para una muestra* aukeratuz (6.1. irudia), 6.2. irudian adierazten den moduan hasiko da kontrastea:



6.1. irudia.



6.2. irudia.

Alde batetik, prozedurak laginaren tamaina, batez besteko, desbideratze tipikoa eta batez bestekoaren errore tipikoa adierazten ditu (6.1. taula).

6.1. taula.

	N	Media	Desviación típ.	Error típ. de la media
Productos	60	108,13	11,048	1,426

Bestalde, 6.2. taulan daude: batetik, kontrasteko estatistikoaren balioa $t = \frac{\bar{x} - \mu_0}{S / \sqrt{n}} = 4,3$, bestetik askatasun-graduak $n - 1 = 59$ eta azkenik p-balio (Sig.) nulua.

6.2. taula.

	Valor de prueba = 102					
	t	gl	Sig. (bilateral)	Diferencia de medias	95% Intervalo de confianza para la diferencia	
					Inferior	Superior
Productos	4,300	59	,000	6,133	3,28	8,99

p-balioa adierazgarritasun-maila baino txikiagoa denez ($0 < 0,5$), errefusatu egingo da $H_0: \mu = 102$ dioen hipotesi nulua, %5eko adierazgarritasun-mailaz.

Gainera, 6.2. taulan %95eko konfiantza-mailako konfiantza-tartearen muturrak daude eta hauen arabera $\mu - 102$ differentzia 3,28 eta 8,99 bitarteko da. Beste hitz batzuetan, %95eko konfiantza-mailaz, asteko batez besteko ekoizpena hurrengo tarteko balio bat izango da:

$$[102 + 3,28, 102 + 8,99] = [105,28, 110,99]$$

102 balioa tartetik kanpo dagoenez, errefusatu egingo da $H_0: \mu = 102$ dioen hipotesi nulua, %5eko adierazgarritasun-mailaz.

6.2.2. adibidea

Demagun A eta B enpresek ekoizten dituzten habeen luzerak banaketa normalekoak eta elkarrekiko askeak direla. Aleatorioki A eta B enpresetan ekoitzitako 12 eta 16 habe hartu dira, eta ondoko luzerak (zentimetrotan) dituzte:

A enpresako habeak: 16; 14,5; 15; 14; 16,5; 15,5; 15; 14,5; 16,5; 16; 15; 14,5.

B enpresako habeak: 14; 15; 16; 14,5; 15,5; 14,5; 16,5; 16; 14,5; 16; 15,5; 14,5; 16; 15,5; 16; 16,5.

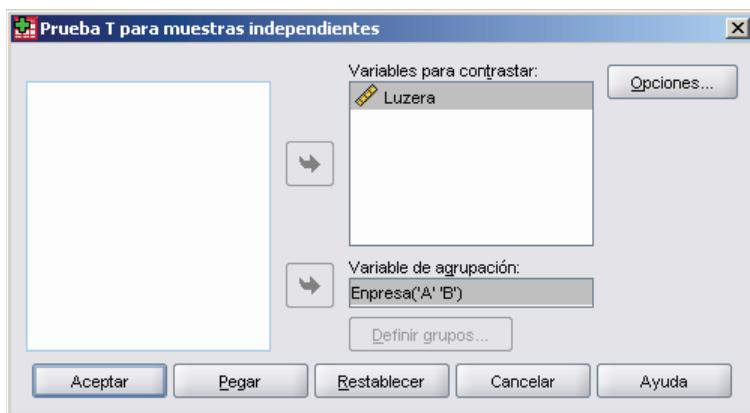
- a) %99ko konfiantza-mailaz, lortu batez besteko luzeren arteko differentzia estimatzeko konfiantza-tartea.
- b) %1eko adierazgarritasun-mailaz, onartuko al zenuke batez besteko luzeren arteko differentziarik ez dagoela dioen hipotesia?

Ebazpena:

Lehenengo datuak sartuko dira SPSS programan eta lagin independenteetarako T proba egingo da, 6.3. irudian adierazten den moduan:

6.3. irudia.

Analizar > Comparar medias > Prueba T para muestras independientes prozedura aukeratuz, 6.4. irudiko koadroa irekiko da:



6.4. irudia.

Ondoren, aleatorioki hartu diren bi laginetako habeen luzeren batez bestekoak eta desbideratze tipikoak kalkulatuko dira (6.3. taula).

6.3. taula.

Procedencia	N	Media	Desviación típ.	Error típ. de la media
Longitud	A	12	15,2500	,83937
	B	16	15,4156	,80864

6.3. taulak adierazten duen moduan, bi laginen batez besteko eta desbideratze tipikoak nahiko antzekoak dira.

Bestalde, 6.4. taulako zenbakiek erakusten dutenez, p-balioak (0,602; 0,605) adierazgarritasun-maila (0,01) baino handiagoak dira. Beraz, %1eko adierazgarritasun-mailaz, onar daiteke A eta B enpresetako habeen luzeren artean differentzia adierazgarririk ez dagoela dioen hipotesia.

6.4. taula.

	Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
	F	Sig.	t	gl	Sig. (bilateral)	Dif. medias	Error típ. de la dif.	95% Intervalo de confianza para la diferencia	
								Inferior	Superior
Longitud Asumido varianzas iguales	,021	,886	-,528	26	,602	-,16563	,31382	-,81070	,47945
No varianzas iguales			-,525	23,348	,605	-,16563	,31556	-,81788	,48663

Bukatzeko, %99ko konfiantza-mailaz, batez bestekoen arteko differentzia estimatzeko konfiantza-tartea hurrengoak dira:

- i) [-0,81070, 0,47945], populazioen bariantzak berdinak direnean.
- ii) [-0,81788, 0,48663], populazioen bariantzak desberdinak direnean.

Bi konfiantza-tarteetan zero zenbakia dagoenez, %99ko konfiantza-mailaz onar daiteke populazioen batez bestekoen arteko berdintza.

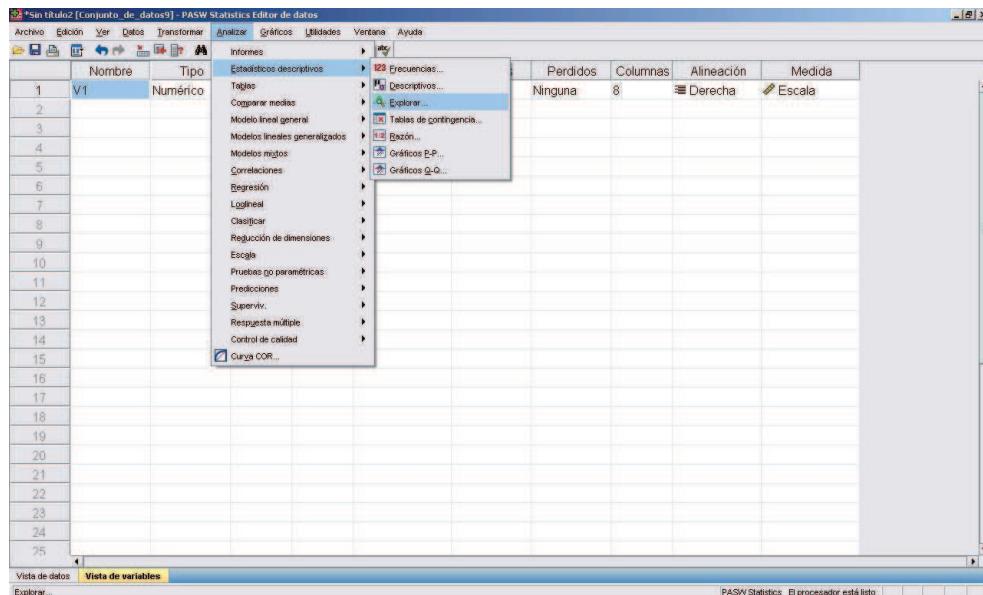
6.2.3. adibidea

%5eko adierazgarritasun-mailaz, frogatu hurrengo datuak banaketa normal batekoak direnetz:

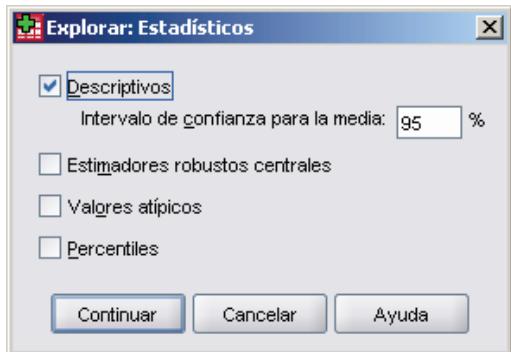
10,2	11,5	11,6	11,2	12,0	9,8	13,0	11,8	11,4	10,6
9,3	10,0	8,9	10,6	11,0	10,0	8,6	10,2	11,4	10,0
9,8	9,5	12,8	10,0	10,6	10,5	10,3	9,9	11,6	11,7
11,5	10,5	11,9	10,8	10,9	11,0	9,2	12,8	11,3	10,8
9,9	9,9	11,0	10,6	10,5	12,0	13,4	13,0	10,5	10,6

Ebazpena:

Datuak *V1* izeneko aldagaien sartuko dira. Gero *Analizar > Estadísticos > Explorar* aukeratuko da (6.5. irudia) eta 6.6. irudiko *Explorar* koadroa irekiko da.

**6.5. irudia.****6.6. irudia.**

6.6. irudiko *Explorar* koadroan *Ambos* delakoa aukeratzen bada, *Estadísticos* eta *Gráficos* botoietan egindako aukeren irteerak erakutsiko dira emaitzen leihon. Horrela, *Estadísticos* azpikoadroan estatistiko deskribatzaileak aukeratu dira (6.7. irudia) eta *Gráficos* azpikoadroan (6.8. irudia) normaltasunaren kontrastea burutzeko grafikoak aukeratu dira.

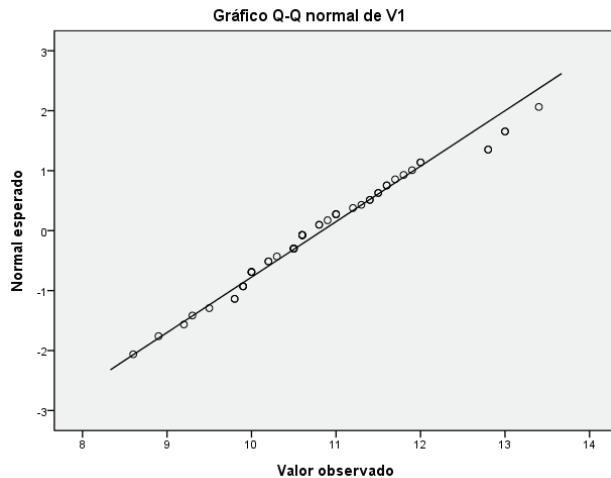


6.7. irudia.



6.8. irudia.

Aginduak exekutatu ondoren, 6.9. irudiko Q-Q grafikoa lortuko da.



6.9. irudia.

Grafiko honetan puntuak ongi egokitzen dira lehenengo koadranteko erdibitzailera, eta ondorioz enuntziatuko aldagaiaren balioak banaketa normalekoak direla esan daiteke, %5eko adierazgarritasun-mailaz.

Halaber, Kolmogorov-Smirnov eta Shapiro-Wilk-en bidezko normaltasun-probetalan (6.5. taula) p-balioak 0,05 adierazgarritasun-maila baino handiagoak direnez, banaketa normala dela dioen hipotesia onartuko da.

6.5. taula.

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
V1	,107	50	,200*	,976	50	,395

a. Corrección de la significación de Lilliefors

*. Este es un límite inferior de la significación verdadera.

6.2.4. adibidea

Fabrikatzaile batek dioenez, bere lantegiko ekoizpenaren %90 kalitate handikoa da. Aleatorioki lantegiko 500 produktu hartu dira eta hauetatik 448 producto kalitate handikoak dira. %2,5eko adierazgarritasun-mailaz, zer esango zenuke fabrikatzailearen baieztapenari buruz?

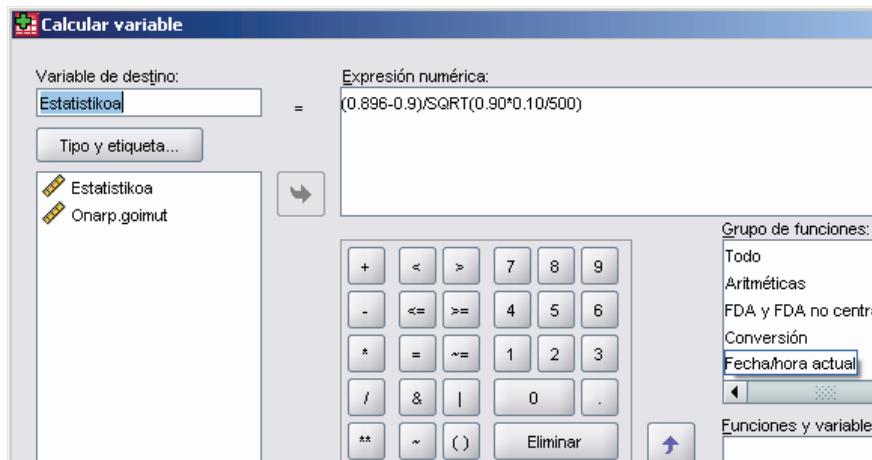
Ebazpena:

Kasu honetan, p = “kalitate handiko produktuen proportzioa” parametroari buruzko kontrastea egindo da. Hartu den lagin aleatorio bakunean kalitate handiko produktuen proportzioa $\hat{p} = \frac{448}{500} = 0,896$ da. Enuntziatuko baieztapenari buruzko

erabakia hartzeko, $H_0: p = 0,90$ hipotesi nulua $H_1: p \neq 0,90$ hipotesi alternatiboarekin kontrasta daiteke, $\alpha = 0,025$ adierazgarritasun-maila hartuz.

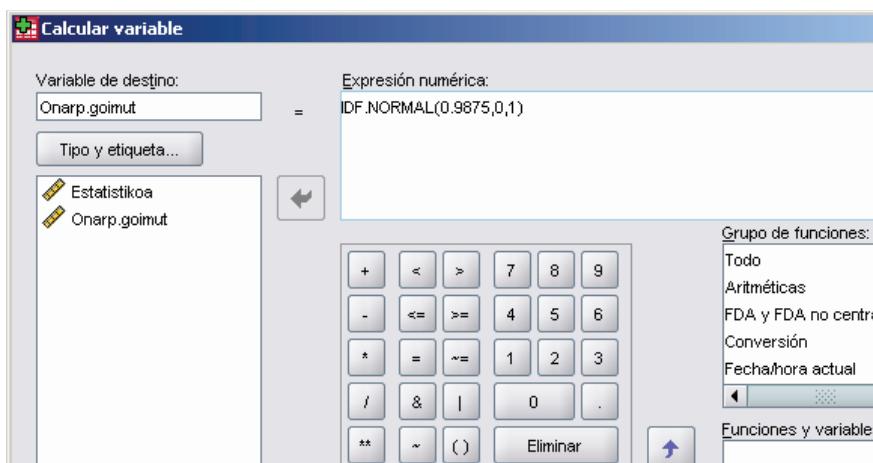
Hipotesi-kontrasteko estatistikoaren balioa *Transformar > Calcular variable* prozeduraren bidez kalkula daiteke (6.10. eta 6.12. irudia):

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0,896 - 0,90}{\sqrt{\frac{0,90 \cdot 0,10}{500}}} = -0,298$$



6.10. irudia.

Bestalde, kasu honetan onarpene-eremua $S_0 = [-z_{0,0125}, z_{0,0125}] = [-2,24, 2,24]$ da (6.11. irudian goi-muturra kalkulatzeko prozedura eta 6.12. irudian balioa adierazidira).



6.11. irudia.

*Sin título1 [Conjunto_de_datos0] - PASW Statistics Editor de datos								
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda								
1 : Estatistikoa - .298142396999972								
1	Estatistikoa	Onarp.goimut	var	var	var	var	var	var
1		- .298	2.24					
2								
3								
4								

6.12. irudia.

Estatistikoaren balioa onarpenean dagoenez, $-0,298 \in S_0$ alegia, $H_0: p = 0,90$ hipotesi nulua onartuko genuke %2,5eko adierazgarritasun-mailaz. Ondorioz, fabrikatzailearen baieztapena onartuko genuke.

6.3. Praktikatzeko ariketak

6.3.1. ariketa

%5eko adierazgarritasun-mailaz, onartuko al zenuke populazioaren batez bestekoa 10 unitatekoa dela dioen hipotesia 6.2.3. adibidean? Lortu batez bestekoa estimatzeko %95eko konfiantza-mailako konfiantza-tartea adibide horretan.

6.3.2. ariketa

A eta B motako piezen gogortasun-indizea neurten duten aldagaiak banaketa normalekoak eta elkarrekiko askeak direla suposatu da. Aleatorioki A motako 16 pieza eta B motako 12 pieza hartu dira, ondoko gogortasun-indizeak dituztelarik:

A motako piezak: 0,28; 0,30; 0,32; 0,29; 0,31; 0,29; 0,33; 0,32; 0,29; 0,32; 0,31; 0,29; 0,32; 0,31; 0,32; 0,33.

B motako piezak: 0,32; 0,29; 0,30; 0,28; 0,33; 0,31; 0,30; 0,29; 0,33; 0,32; 0,30; 0,29.

- a) %98ko konfiantza-mailaz, lortu batez besteko gogortasun-indizeen arteko differentzia estimatzeko konfiantza-tartea.
- b) %2ko adierazgarritasun-mailaz, ba al dago differentzia adierazgaririk batez besteko gogortasun-indizeen artean?

6.3.3. ariketa

%1eko adierazgarritasun-mailaz, banaketa normalekoak al dira 6.2.1. adibideko datuak?

6.3.4. ariketa

Burdinazko habeak egiten dituzten lantegi batean aleatorioki 250 habe hartu dira eta horietatik 75 habek erdoiltzeko joera erakutsi dute. Erdoiltzearen aurkako tratamendu bat aplikatu ondoren, lantegian aleatorioki beste 250 burdinazko habe hartu dira eta horietatik 25 habek erdoiltzeko joera erakutsi dute. %1eko adierazgarritasun-mailaz, onargarria al da “erdoiltzeko joera erakusten duten habeen proportzioa txikiagoa da tratamendua jaso duten habeen kasuan” dioen hipotesia?

Sinboloak eta laburdurak

Sinboloa	Esanahia
Q_i	i. ordenako kuartila
S^2	Laginaren kuasibariantza
x_i	Balioak, behaketak
D_j	j. ordenako dezila
s	Laginaren desbideratze tipikoa edo desbideratze estandarra
σ	Populazioaren desbideratze tipikoa edo desbideratze estandarra
ε	Errorea, estimazio-errorea
S_0	Onarpen-eremua
F	Snedecor-en F estatistikoa
t	Student-en t estatistikoa
gl	Askatasun-graduak
H_1	Hipotesi alternatiboa
H_0	Hipotesi nula
$I_\theta^{1-\alpha}$	θ estimatzeko $1-\alpha$ konfiantza-mailako konfiantza-tartea
\bar{x}	Laginaren batez bestekoa
μ	Populazioaren batez bestekoa
$1 - \alpha$	Konfiantza-maila
n	Laginaren tamaina, datu kopurua, aldagai kopurua
θ	Populazioaren parametroa
P_k	k. ordenako pertzentila
P	Probabilitatea
α	Adierazgarritasun-maila
\hat{p}	Laginaren proportzioa
p	Populazioaren proportzioa
z	Banaketa normal estandarreko balioa
$Var(X)$	X aldagaiaren bariantza
s^2	Laginaren bariantza

σ^2 Populazioaren bariantza

Laburdura	Esanahia
<i>Binom(n,p)</i>	n eta p parametroetako banaketa binomiala
<i>CDF</i>	Cumulative density function
<i>Exp(β)</i>	β parametroko banaketa esponentziala
<i>IDF</i>	Inverse density function
<i>N(0, 1)</i>	Banaketa normal tipiko edo estandarra
<i>N(μ, σ)</i>	μ eta σ parametroetako banaketa normala
<i>PDF</i>	Probability density function
<i>Poisson(λ)</i>	λ parametroko Poisson-en banaketa
<i>Sig</i>	p-balioa (SPSSren notazioan)
<i>SQRT</i>	Erro karratua
<i>Uniform(a,b)</i>	Banaketa uniformea [a, b] tartean
χ^2_v	v askatasun-graduko khi karratu banaketa
t_v	v askatasun-graduko Student-en t banaketa
T	T-proba
F_{v_1, v_2}	v_1 askatasun-gradu zenbakitzaillean eta v_2 askatasun-gradu zatitzaillean dituen F banaketa

Bibliografía

- Agirre, E.** (2010). *Estatistikaren oinarriak. Ariketak (2. argitalpena)*. Udako Euskal Unibertsitatea, Bilbo.
- Barron, L.J.R., Agirre, E.** (2011). *Guía de Métodos Estadísticos en Calidad y Seguridad Alimentaria*. Barron,L.J.R., Agirre, E. Vitoria-Gasteiz.
- Etxeberria, J.** (2004). *Estatistika eta SPSS*. Elhuyar, Usurbil.
- Etxeberria, J.** (1998). *Estatistika eta praktika SPSSWIN erabiliz*. Elhuyar, Usurbil.
- Ferrán, M.** (1996). *SPSS para Windows. Programación y análisis estadístico*. McGraw-Hill, Madrid.
- Isasi, X.** (2010). *Erregresio lineal, bariantza-analisiak eta hipotesi-testak*. Udako Euskal Unibertsitatea, Bilbo.
- Johnson, D.E.** (2000). *Métodos multivariados aplicados al análisis de datos*. Thomson, Madrid.
- Johnson, R.A.** (1997). *Probabilidad y Estadística para Ingenieros de Miller y Freund*. Prentice Hall Hispanoamericana, México.
- Martín, Q., Cabero, M.T., Del Rosario de Paz, Y.** (2007). *Tratamiento estadístico de datos con SPSS. Prácticas resueltas y comentadas*. Paraninfo, S.A.
- Mendenhall, W, Sincich, T.** (1997). *Probabilidad y Estadística para Ingeniería y Ciencias*.Prentice Hall Hispanoamericana, México.
- Pérez López, C.** (2005) *Métodos estadísticos avanzados con SPSS*. Thomson, Madrid.
- Pérez López, C.** (2009). *Técnicas de Análisis de Datos con SPSS15*. Pearson Educación España, Madrid.

ISBN 978-84-615-7798-9



9 788461 577989