



Reporte de Práctica 5: Pruebas estadísticas

En ésta práctica se realizarán 3 análisis estadísticos a los datos del cine. Primero realizaremos histogramas de la información para observar su comportamiento si es normalizado o no.

Como se harán gráficas de los datos, se instaló la librería Plotly de python3 para hacer esta tarea, debido a que es una librería de pago, se creó una cuenta gratuita en la cual estamos limitados a 25 gráficas.

Se usaron las siguientes intrucciones para configurar el Plotly.

In [16]: `!pip3 install plotly`

```
Requirement already satisfied: plotly in /usr/local/lib/python3.6/dist-packages (3.6.1)
Requirement already satisfied: nbformat>=4.2 in /usr/local/lib/python3.6/dist-packages (from plotly) (4.4.0)
Requirement already satisfied: pytz in /usr/lib/python3/dist-packages (from plotly) (2018.3)
Requirement already satisfied: six in /home/samataroukami/.local/lib/python3.6/site-packages (from plotly) (1.11.0)
Requirement already satisfied: decorator>=4.0.6 in /usr/local/lib/python3.6/dist-packages (from plotly) (4.3.0)
Requirement already satisfied: retrying>=1.3.3 in /usr/local/lib/python3.6/dist-packages (from plotly) (1.3.3)
Requirement already satisfied: requests in /home/samataroukami/.local/lib/python3.6/site-packages (from plotly) (2.20.1)
Requirement already satisfied: ipython-genutils in /usr/local/lib/python3.6/dist-packages (from nbformat>=4.2->plotly) (0.2.0)
Requirement already satisfied: jsonschema!=2.5.0,>=2.4 in /usr/local/lib/python3.6/dist-packages (from nbformat>=4.2->plotly) (2.6.0)
Requirement already satisfied: traitlets>=4.1 in /usr/local/lib/python3.6/dist-packages (from nbformat>=4.2->plotly) (4.3.2)
Requirement already satisfied: jupyter-core in /usr/local/lib/python3.6/dist-packages (from nbformat>=4.2->plotly) (4.4.0)
Requirement already satisfied: idna<2.8,>=2.5 in /home/samataroukami/.local/lib/python3.6/site-packages (from requests->plotly) (2.7)
Requirement already satisfied: certifi>=2017.4.17 in /home/samataroukami/.local/lib/python3.6/site-packages (from requests->plotly) (2018.10.15)
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in /home/samataroukami/.local/lib/python3.6/site-packages (from requests->plotly) (3.0.4)
Requirement already satisfied: urllib3<1.25,>=1.21.1 in /home/samataroukami/.local/lib/python3.6/site-packages (from requests->plotly) (1.24.1)
```

In [21]: `!pip install --user --upgrade pip`

```
Cache entry deserialization failed, entry ignored
Collecting pip
  Cache entry deserialization failed, entry ignored
  Downloading https://files.pythonhosted.org/packages/d8/f3/413bab4ff08e1fc4828dfc59996d721917df8e8583ea85385d51125dceff/pip-19.0.3-py2.py3-none-any.whl (1.4MB)
    100% |████████████████████████████████████████| 1.4MB 581kB/s ta 0:00:01
Installing collected packages: pip
  Found existing installation: pip 9.0.3
  Uninstalling pip-9.0.3:
    Successfully uninstalled pip-9.0.3
Successfully installed pip-19.0.3
```

```
In [41]: !python -m pip install plotly --user
```

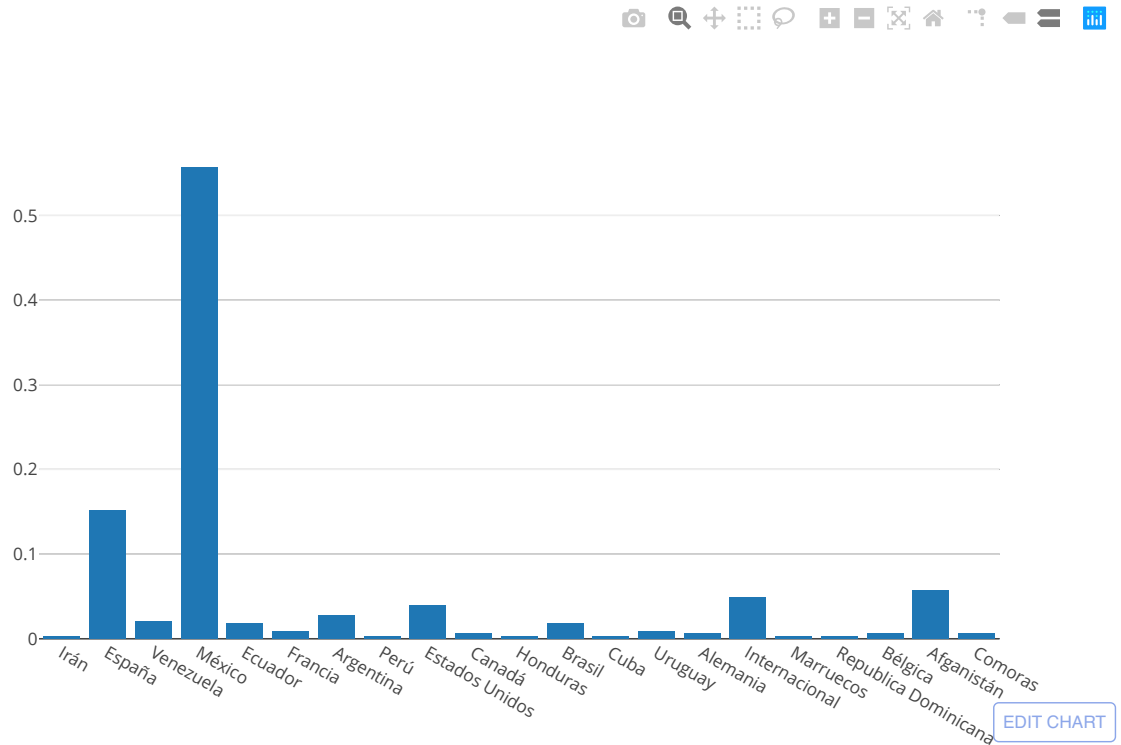
```
Collecting plotly
  Using cached https://files.pythonhosted.org/packages/fd/db/003b5cfbc710f4d4982440451185b952269e4080a57ae7e760a2ceb8ce0c/plotly-3.6.1-py2.py3-none-any.whl
Requirement already satisfied: six in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (1.11.0)
Requirement already satisfied: decorator>=4.0.6 in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (4.3.2)
Requirement already satisfied: requests in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (2.20.1)
Requirement already satisfied: nbformat>=4.2 in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (4.4.0)
Collecting retrying>=1.3.3 (from plotly)
  Using cached https://files.pythonhosted.org/packages/44/ef/beae4b4ef80902f22e3af073397f079c96969c69b2c7d52a57ea9ae61c9d/retrying-1.3.3.tar.gz
Requirement already satisfied: pytz in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (2018.9)
Requirement already satisfied: idna<2.8,>=2.5 in /home/samataroukami/.local/lib/python2.7/site-packages (from requests->plotly) (2.7)
Requirement already satisfied: urllib3<1.25,>=1.21.1 in /home/samataroukami/.local/lib/python2.7/site-packages (from requests->plotly) (1.24.1)
Requirement already satisfied: certifi>=2017.4.17 in /home/samataroukami/.local/lib/python2.7/site-packages (from requests->plotly) (2018.10.15)
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in /home/samataroukami/.local/lib/python2.7/site-packages (from requests->plotly) (3.0.4)
Requirement already satisfied: traitlets>=4.1 in /home/samataroukami/.local/lib/python2.7/site-packages (from nbformat>=4.2->plotly) (4.3.2)
Requirement already satisfied: jsonschema!=2.5.0,>=2.4 in /home/samataroukami/.local/lib/python2.7/site-packages (from nbformat>=4.2->plotly) (3.0.0)
Requirement already satisfied: ipython-genutils in /home/samataroukami/.local/lib/python2.7/site-packages (from nbformat>=4.2->plotly) (0.2.0)
Requirement already satisfied: jupyter-core in /home/samataroukami/.local/lib/python2.7/site-packages (from nbformat>=4.2->plotly) (4.4.0)
Requirement already satisfied: enum34; python_version == "2.7" in /home/samataroukami/.local/lib/python2.7/site-packages (from traitlets>=4.1->nbformat>=4.2->plotly) (1.1.6)
Requirement already satisfied: pyparsing>=0.14.0 in /home/samataroukami/.local/lib/python2.7/site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.2->plotly) (0.14.1)
Requirement already satisfied: functools32; python_version < "3" in /home/samataroukami/.local/lib/python2.7/site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.2->plotly) (3.2.3.post2)
Requirement already satisfied: attrs>=17.4.0 in /home/samataroukami/.local/lib/python2.7/site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.2->plotly) (18.2.0)
Requirement already satisfied: setuptools in /home/samataroukami/.local/lib/python2.7/site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.2->plotly) (40.6.2)
Building wheels for collected packages: retrying
  Running setup.py bdist_wheel for retrying ... done
  Stored in directory: /home/samataroukami/.cache/pip/wheels/d7/a9/33/acc7b709e2a35caa7d4cae442f6fe6fbf2c43f80823d46460c
Successfully built retrying
Installing collected packages: retrying, plotly
Successfully installed plotly-3.6.1 retrying-1.3.3
You are using pip version 18.1, however version 19.0.3 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
```

```
Requirement already up-to-date: plotly in /home/samataroukami/.local/lib/python2.7/site-packages (3.6.1)
Collecting pip
  Using cached https://files.pythonhosted.org/packages/d8/f3/413bab4ff08e1fc4828dfc59996d721917df8e8583ea85385d51125dceff/pip-19.0.3-py2.py3-none-any.whl
Requirement already satisfied, skipping upgrade: six in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (1.11.0)
Requirement already satisfied, skipping upgrade: decorator>=4.0.6 in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (4.3.2)
Requirement already satisfied, skipping upgrade: requests in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (2.20.1)
Requirement already satisfied, skipping upgrade: nbformat>=4.2 in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (4.4.0)
Requirement already satisfied, skipping upgrade: retrying>=1.3.3 in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (1.3.3)
Requirement already satisfied, skipping upgrade: pytz in /home/samataroukami/.local/lib/python2.7/site-packages (from plotly) (2018.9)
Requirement already satisfied, skipping upgrade: idna<2.8,>=2.5 in /home/samataroukami/.local/lib/python2.7/site-packages (from requests->plotly) (2.7)
Requirement already satisfied, skipping upgrade: urllib3<1.25,>=1.21.1 in /home/samataroukami/.local/lib/python2.7/site-packages (from requests->plotly) (1.24.1)
Requirement already satisfied, skipping upgrade: certifi>=2017.4.17 in /home/samataroukami/.local/lib/python2.7/site-packages (from requests->plotly) (2018.10.15)
Requirement already satisfied, skipping upgrade: chardet<3.1.0,>=3.0.2 in /home/samataroukami/.local/lib/python2.7/site-packages (from requests->plotly) (3.0.4)
Requirement already satisfied, skipping upgrade: traitlets>=4.1 in /home/samataroukami/.local/lib/python2.7/site-packages (from nbformat>=4.2->plotly) (4.3.2)
Requirement already satisfied, skipping upgrade: jsonschema!=2.5.0,>=2.4 in /home/samataroukami/.local/lib/python2.7/site-packages (from nbformat>=4.2->plotly) (3.0.0)
Requirement already satisfied, skipping upgrade: ipython-genutils in /home/samataroukami/.local/lib/python2.7/site-packages (from nbformat>=4.2->plotly) (0.2.0)
Requirement already satisfied, skipping upgrade: jupyter-core in /home/samataroukami/.local/lib/python2.7/site-packages (from nbformat>=4.2->plotly) (4.4.0)
Requirement already satisfied, skipping upgrade: enum34; python_version == "2.7" in /home/samataroukami/.local/lib/python2.7/site-packages (from traitlets>=4.1->nbformat>=4.2->plotly) (1.1.6)
Requirement already satisfied, skipping upgrade: pyrsistent>=0.14.0 in /home/samataroukami/.local/lib/python2.7/site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.2->plotly) (0.14.11)
Requirement already satisfied, skipping upgrade: functools32; python_version < "3" in /home/samataroukami/.local/lib/python2.7/site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.2->plotly) (3.2.3.post2)
Requirement already satisfied, skipping upgrade: attrs>=17.4.0 in /home/samataroukami/.local/lib/python2.7/site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.2->plotly) (18.2.0)
Requirement already satisfied, skipping upgrade: setuptools in /home/samataroukami/.local/lib/python2.7/site-packages (from jsonschema!=2.5.0,>=2.4->nbformat>=4.2->plotly) (40.6.2)
Installing collected packages: pip
  Found existing installation: pip 18.1
    Uninstalling pip-18.1:
      Successfully uninstalled pip-18.1
Successfully installed pip-19.0.3
```

```
In [43]: import plotly
plotly.tools.set_credentials_file(username='SamatarouKami', api key='X00Ga3VP5G6DWcfBivYa')
```

```
In [2]: import plotly.plotly as py
import plotly.graph_objs as go
import pandas as pd
cine = pd.read_csv('https://raw.githubusercontent.com/SamatarouKami/CIENCIA_DE_DATOS/master/datosLimpiosCine2.csv', sep='\\')
ext = cine.Pais[cine.Pais!='Colombia']
print((py.iplot({'data': [go.Histogram(x = ext, histnorm='probability')]}), filename='extranjeros')).
embed_code)
```

```
<iframe id="igraph" scrolling="no" style="border:none;" seamless="seamless" src="https://plot.ly/~SamatarouKami/2.embed" height="525px" width="100%"></iframe>
```



Se supone que aqui arriba hay un gráfico. Así generamos el histograma de la cantidad de paises extranjeros que participaron.

```
In [5]: !python -m pip install statsmodels --user
```

DEPRECATION: Python 2.7 will reach the end of its life on January 1st, 2020. Please upgrade your Python as Python 2.7 won't be maintained after that date. A future version of pip will drop support for Python 2.7.

Collecting statsmodels

Downloading https://files.pythonhosted.org/packages/ed/56/ald32debdbaba5cc7986aeceb79e33d74ae3394eabfd008b0113368b91981/statsmodels-0.9.0-cp27-cp27mu-manylinux1_x86_64.whl (7.4MB)

100% |██| 7.4MB 1.1MB/s ta 0:00:01

Requirement already satisfied: pandas in /home/samataroukami/.local/lib/python2.7/site-packages (from statsmodels) (0.24.1)

Collecting patsy (from statsmodels)

Using cached <https://files.pythonhosted.org/packages/ea/0c/5f61f1a3d4385d6bf83b83ea495068857ff8dfb89e74824c6e9eb63286d8/patsy-0.5.1-py2.py3-none-any.whl>

Requirement already satisfied: numpy>=1.12.0 in /home/samataroukami/.local/lib/python2.7/site-packages (from pandas->statsmodels) (1.16.1)

Requirement already satisfied: pytz>=2011k in /home/samataroukami/.local/lib/python2.7/site-packages (from pandas->statsmodels) (2018.9)

Requirement already satisfied: python-dateutil>=2.5.0 in /home/samataroukami/.local/lib/python2.7/site-packages (from pandas->statsmodels) (2.8.0)

Requirement already satisfied: six in /home/samataroukami/.local/lib/python2.7/site-packages (from patsy->statsmodels) (1.11.0)

Installing collected packages: patsy, statsmodels

Successfully installed patsy-0.5.1 statsmodels-0.9.0

```
In [7]: import pandas as pd
from statsmodels.graphics.gofplots import qqplot
import matplotlib as plt

cine = pd.read_csv('https://raw.githubusercontent.com/SamatarouKami/CIENCIA_DE_DATOS/master/datosLimpiosCine2.csv', sep='\\')
ext = cine.Pais[cine.Pais!='Colombia']

py.ipplot({'data': [go.Histogram(x = ext, histnorm='probability')]}, filename='extranjeros')

f = qqplot(ext, line='s')
plt.save(f, "qq_cfep.png")
```

```
-----
TypeError                                 Traceback (most recent call last)
<ipython-input-7-3e0b50d9d610> in <module>()
      9
     10
--> 11 f = qqplot(ext, line='s')
     12 plt.save(f, "qq_cfep.png")

/home/samataroukami/.local/lib/python2.7/site-packages/statsmodels/graphics/gofplots.pyc in qqplot
(data, dist, distargs, a, loc, scale, fit, line, ax, **plotkwargs)
    504     """
    505     probplot = ProbPlot(data, dist=dist, distargs=distargs,
--> 506                        fit=fit, a=a, loc=loc, scale=scale)
    507     fig = probplot.qqplot(ax=ax, line=line, **plotkwargs)
    508     return fig

/home/samataroukami/.local/lib/python2.7/site-packages/statsmodels/graphics/gofplots.pyc in __init__
__(self, data, dist, fit, distargs, a, loc, scale)
    135         dist = getattr(stats, dist)
    136
--> 137         self.fit_params = dist.fit(data)
    138         if fit:
    139             self.loc = self.fit_params[-2]

/home/samataroukami/.local/lib/python2.7/site-packages/scipy/stats/_continuous_distns.pyc in fit(s
elf, data, **kwargs)
    266
    267         if floc is None:
--> 268             loc = data.mean()
    269         else:
    270             loc = floc

/home/samataroukami/.local/lib/python2.7/site-packages/numpy/core/_methods.pyc in _mean(a, axis, d
type, out, keepdims)
    73         is_float16_result = True
    74
--> 75         ret = umr_sum(arr, axis, dtype, out, keepdims)
    76         if isinstance(ret, mu.ndarray):
    77             ret = um.true_divide(

TypeError: cannot concatenate 'str' and 'float' objects
```

Conclusiones

No pude hacer que se vieran en la práctica los gráficos, por lo menos no se ven en mi computadora, espero que mañana se vean, los análisis estadísticos requieren que de alguna forma yo convierta mis datos a números ya que manejo muchas cadenas de caracteres y eso no me sirve para hacer análisis estadísticos, tal vez el objetivo de esta práctica era encontrar o buscar la manera de hacer eso, pero no encontré forma de hacerlo, solo puedo hacer conteos de la información. Tengo que categorizar.

--04 de Marzo 2019-- Luis Angel Gutierrez Rodriguez [1484412](tel:1484412) (tel:1484412)