TEST DI IPOTESI

In seguito alla rassegna dei parametri fondamentali che caraterizzano gli esopianeti, si vuole sottoporre tali elementi a uno studio che ci permetta di dedurre un modello che possa predire la temperatura la loro temperatura. A tale fine si vogliono supportare attraverso i test d’ipotesi due asserzioni:

1. La temperatura non dipende dal numero di stelle del sistema di uci fa parte
2. I metodi di scoperta hanno generato un forte bias nel pool dei dati

Si procede, dunque, suddividendo in categorie minori le temperature:

* Per **numero di stelle**:
  + Chart, histogram

    Description automatically generatedTemperature dei pianeti con una **singola** stella:
    - **T\_st1= 692 dati**

**Classi: 1+log\_2(692)=10.43**

****

Chart, histogram

Description automatically generated

* + Temperature dei pianeti con più di **due** stelle
    - **T\_st234=112 dati**

**Classi: 1+log\_2(112)=8.8**

A picture containing text

Description automatically generated

* Per **metodo di scoperta:**
  + Chart, histogram

    Description automatically generatedImaging**:**
    - **T\_ima= 27dati**
    - **Text

      Description automatically generated with medium confidenceClassi=1+log\_2(27)=5.75**

****

* + Radial Velocity:
    - **Chart, histogram

      Description automatically generatedT\_ rv=55 dati**
    - **Classi=1+log\_2(55)=6.78**

****

****

* + Transit:
    - **T\_trans: 715 dati**
    - **Chart, histogram

      Description automatically generatedClassi=1+log\_2(715)=10.48**

****

****

**Analisi normalità campioni:**

In seguito alla suddivisione delle temperature in base alle categorie descritte, si studia la gaussianità delle relative distribuzioni.

**Temperatura 1 stella Temperatura 2+ stelle**

Chart, line chart

Description automatically generated**Chart, line chart

Description automatically generated**

Text

Description automatically generated

**Text

Description automatically generated**

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generatedChart, line chart

Description automatically generatedText

Description automatically generated **Temperatura Radial Velocity Temperatura Transit Temperatura Imaging**

Text

Description automatically generated with low confidenceText

Description automatically generated with medium confidence

Nonostante la numerosità dei campioni abbia influenzato la potenza del test di Shapiro-Wilks, rendendolo molto suscettibile alle minime variazioni dalla retta dei quantili, i qq plot ci spingono ad affermare l’ipotesi alternativa di non gaussinità. Caso d’eccezione spetta al T\_ima che non una numerosità bassa di n=27, presenta un p-value di 0.1554. In questo caso non vi sono forte evidenze per rifiutare l’ipotesi nulla, dunque si accetta la normalità della sua distribuzione.

Per il teorema del limite centrale, poiché campioni numerosi (n>30), possiamo assumere che le medie campionarie abbiano distribuzione normale.

**Numero di stelle:**

* Lo scopo del test è di supportare l’idea che la temperatura non sia correlata dal numero di stelle appartenenti al sistema. A tal fine si effettua un Z-test su due campioni numerosi di tipo bilaterale.

con

**“La temperatura dei pianeti con una singola stella è maggiore dei pianeti con più stelle?”**

**Graphical user interface, text, email

Description automatically generated**

Il test presenta un p-value di 0.1966. Non vi sono evidenze forti per rifiutare l’ipotesi nulla.   
Si conclude che la temperatura media dei pianeti non è dipendente dal numero di stelle appartenenti al sistema.

**Metodo di scoperta:**

* Si procede ora con il definire un secondo Z-test di tipo bilaterale che verifichi la differenza di temperatura media rilevata in base alla metodologia utilizzata. Il confronto avviene fra i metodi di ***Transit*** *e* ***Radial Velocity***, sulle quali è applicabile il TLC.

**Graphical user interface, text, application, email

Description automatically generated“I metodi di rilevazione influenzano con un forte bias la temperatura media rilevata?”**

Il p-value estremamente basso ci permette d iavere forte evidenze contro l’ipotesi nulla. Dunque, si assume vera l’ipotesi alternativa per cui le medie non coincidono.

* Graphical user interface, text

  Description automatically generatedIl successivo Z-test tra **Radial Velocity** e **Imaging** assume la normalità della media campionaria dei dati tratti dall’Imaging sebbene essa presenti una numerosità piuttosto borderline **n=27<30**. Tale assunzione è validà poiché si è verificato con un test di Shapiro-Wilks che la sua normalità non è rifiutabile. Lo stesso discorso vale per la varianza ignota, stimabile con una buona precisione con

=

Il risultato che si ottiene, come si poteva già presuporre, presenta un p-value estremamente basso, dando forti evidenze per rifiutare l’ipotesi nullo e accettare quello alternativo.

* In ques’ultimo test fra **Transit** e **Imaging** si tengono tutte le supposizioni fatte nel punto precedente.

Graphical user interface, text, email

Description automatically generated

Il test presenta p-value basso. Si rifiuta l’ipotesi nulla e si accetta quella alternativa.

In conclusione si può affermare che i metodi di segnalazione dei pianeti varia fortemente in dipendenza della temperatura, per cui mel pool vi è presenta un forte bias in favore del metodo più utilizzato(Transit).

**Temperatura sopravvivenza rover:**

Un ruolo fondamentale nello studio dei pianeti extraterrestri viene ricoperto da dei piccoli ma sofisticati robot, capaci di operare sotto condizioni estreme per fornire un grande quantitativo di dati necessari allo studio dell’universo che ci circonda: **i rover**.

Si vuole ora studiare la probabilità che un rover possa sopravvivere sugli esopianeti. A tal fine, si considera come modello uno dei più famosi se non più resistenti rover, in grado di resistere fino a temperature estreme di **1300°C** : Perseverance. Per la natura di **pl\_eqt**, parametro in nostro possesso, si suppone che tutti i pianeti siano dei corpi neri a temperatura costante.

Procediamo dunque effettuando un test di proporzione del campione. La numerosità del campione n=804>30, ci permette di trovare uno stimatore **p̂** preciso del campione Bernoulliano. Si assume che ciascuna v.a. i.i.d assume valore **“1”** per o **“0”** per .

**“ La probabilità che Perseverance resista alla temperatura del pianeta è inferiore al 79%”**

***Text, letter

Description automatically generated***

Dato il p-value=0.79>0.05, non vi sono abbastanza evidenze per rifiutare l’ipotesi nulla e accettare l’ipotesi alternativa. Dunque si conclude che la probabilità che il rover possa sopravvivere alla temperatura di un pianeta qualsiasi è maggiore del 79%.

Costruiamo un intervallo di confidenza al 95% per ottenere un range di valori possibili per p.  
Si utilizza il valore stimato

Text

Description automatically generatedper costruire un IC(95%)

Prendendo come riferimento il risultato ottenuto da prop.test, costruiamo l’intervallo

Grazie al risultato del test, possiamo restringere l’intervallo a