MATH 425 12/7/2022

Review

χ^2 test $\displaystyle\sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$ χ^2_{n-1}

Suppose that in a bin there are 10%, 10%, 20%, 20%, 40%, articles of kind 1, 2, 3, 4, 5. After pulling out a sample of 100 article, I got

15, 5, 25, 25, 30

Can I say with 95% confidence that the percentages were wrong?

$$\frac{(15-10)^2}{10} + \frac{(5-10)^2}{10} + \frac{(25-20)^2}{20} + \frac{(25-20)^2}{20} + \frac{(30-40)^2}{40}$$

$= 2.5 + 2.5 + 1.25 + 1.25 + 2.5 = 10 > 9.488$ YES, WE KNOW 95% CONFIDENCE THAT THE % ARE WRONG t-TEST

$$\frac{\bar{X} - \mu}{\sqrt{\sum \frac{(X_i - \bar{X})^2}{n(n-1)}}}$$

<u>Example</u>: 4 samples whose quality scores were 25, 25, 25, 33. The average quality score is 30. Can we say with 95% confidence (a) That the sample is worse than average

(b) The sample is not average.

$4\overline{)108} = 27 = \bar{X}$ $\quad \mu = 30 \quad \frac{\bar{X} - \mu}{\sqrt{4}} = \frac{-3}{2} = -1.5$ 3 d.f.

$$\frac{(25-27)^2 + (25-27)^2 + (25-27)^2 + (33-27)^2}{4 \cdot 3} = \frac{4+4+4+36}{12} = 4$$

1-tailed critical value 3 d.f: 2.353 ← below average

3.182 ← not average

2-tailed

We got 1.5 in both, so we can say nothing

Z test

Suppose we toss a coin 100 times. What is the probability it comes up H > 55 times if it is fair?

> 55.5 ← discrete variable correction (average of 55 and 56)

$\mu = 50$

$\sigma = \sqrt{n \cdot p(1-p)} = \sqrt{100 \cdot \frac{1}{4}} = 5$

$Z > \frac{55.5 - 50}{5} = 1.1 = 1 - P(Z \leq 1.1) = 1 - 0.8643 = \underline{\underline{0.1357}}$

Γ-distribution $\Gamma_{\lambda,\alpha}$ ($\Gamma_{\alpha,\lambda}$ in notes)

hazard rate r in neg. distribution

$$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x} (\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & \text{else} \end{cases}$$

---

20 calls per hour into a center. The distribution of the time in minutes when the 10th caller calls.

$\Gamma_{\lambda,\alpha}$ $\lambda = 1/3$ $\alpha = 10$ $20/60 = 1/3$

density $\left(\frac{1}{3} e^{-x/3} (\frac{1}{3}x)^9\right)/9!$ $x \geq 0$ 0 else

Normal approx:

$\mu = 30$

$\sigma = 3\sqrt{10}$

$Z_{30, 3\sqrt{10}}$

$\dfrac{\Gamma_{1/3,10} - 30}{3\sqrt{10}} \sim Z$

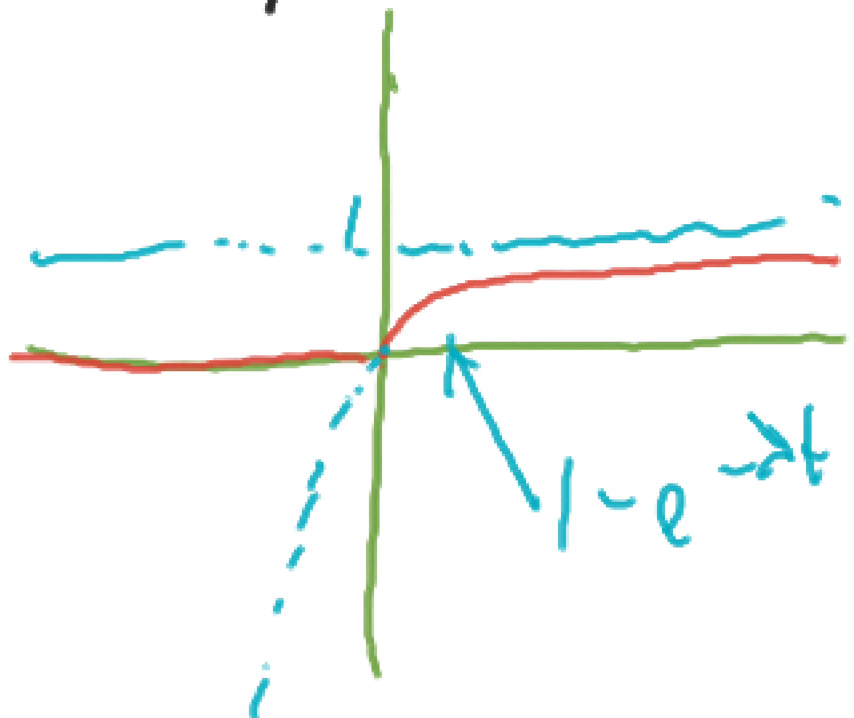
Exponential distribution $= \Gamma_{\lambda,1}$ half-life: $t = \frac{\ln 2}{\lambda}$

Example: If a substance has half-life 7 years, how much of the substance will be left after 10 years?

$$\left(\frac{1}{2}\right)^{10/7}$$

The cumulative distribution for the exponential:

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & \text{else} \end{cases}$$



X, Y jointly distributed random variables.

$$P(X=1 \ \& \ Y=1) = 0.5$$

$$P(X=2 \ \& \ Y=1) = 0.3$$

$$P(X=2 \ \& \ Y=0) = 0.2$$

Are X, Y independent?

$$P(X=s \ \& \ Y=t) = P(X=s) \cdot P(Y=t)$$

$s=t=1$

$$\underbrace{P(X=1 \ \& \ Y=1)}_{0.5} \neq \underbrace{P(X=1)}_{0.5} \underbrace{P(Y=1)}_{0.8}$$

NO

correlation coeff:

$E(X) = 1 \cdot 0.5 + 2 \cdot 0.5 = 1.5$

$E(X^2) = 1 \cdot 0.5 + 4 \cdot 0.5 = 2.5$

$var(X) = 0.25 \quad \sigma(X) = 0.5$

$E(Y) = 1 \cdot 0.8 = 0.8$

$E(Y^2) = 1 \cdot 0.8 = 0.8$

$var(Y) = 0.8 - 0.64 = 0.16$

$\sigma(Y) = 0.4$

$E(XY) = 1 \cdot 0.5 + 2 \cdot 0.3 = 1.1$

$cov(X,Y) = E(XY) - E(X)E(Y)$

$= 1.1 - 1.5 \cdot 0.8 = -0.1$

$$\rho(X,Y) = \frac{-0.1}{0.5 \cdot 0.4} = -0.5$$

X is a continuous random variable with density

$$f(x) = \begin{cases} 4x^3 & 0 \le x \le 1 \\ 0 & \text{else.} \end{cases}$$

What is the density $g(y)$ of $Y = X^2$?

$y = x^2$

$dy = 2x\,dx$

$$f(x)\,dx = g(y)\,dy$$

$$4x^3 \cancel{dx} = g(y)\,2x\cancel{dx}$$

$$2x^2 = g(y)$$

$0 \le x \le 1$

$x = \sqrt{y}$

self-contained, must be in the same variable on both sides!

$$\boxed{\text{Answer: } g(y) = \begin{cases} 2y & 0 \le y \le 1 \\ 0 & \text{else} \end{cases}}$$