

BÀI KIỂM TRA NĂNG LỰC DATA SCIENCE

PHẦN XÁC SUẤT:

1. Một người đàn ông tên X bị cảnh sát tình nghi là hung thủ trong một vụ án mạng. Sau một khoảng thời gian điều tra, cảnh sát thu thập được 2 thông tin như sau:
 - Hiện trường có 3 người: 1 người là nạn nhân, 1 người là X và 1 người cảnh sát không biết là ai và 1 trong 2 người là hung thủ,
 - Thủ phạm là đàn ông.

=> Xác suất hung thủ là X chiếm bao nhiêu phần trăm?
2. Giá xăng dao động rất mạnh trong những năm gần đây. Giả sử giá xăng năm 2019 (tính theo VNĐ) trong đoạn $[10,000; 300,000]$ cho 1 lít (Dưới 10,000đ thì bán xăng sẽ không có lãi và trên 300,000đ thì ta nên dùng loại năng lượng khác tiết kiệm hơn). Giả sử ra có sự kiện B_x ($x = 0, 1, 2, \dots, 9$), có thể coi B_x là họ độc lập các sự kiện được không khi $B_x =$ “Năm 202 x , giá xăng tăng lên ít nhất 50% trong năm đó với giá xăng tính từ năm 2019”.
Giải thích nguyên nhân vì sao?
3. Giả sử Y là một biến ngẫu nhiên liên tục với hàm mật độ $P_{Y(x)} = b \cdot \sin x$ khi $x \in [0; \pi]$ và $P_{Y(x)} = 0$ tại các điểm còn lại:
 - o Tính b ,
 - o Tính $E(Y)$.

PHẦN THỐNG KÊ:

4. Giả sử chiều cao nữ giới Việt Nam tuân theo phân phối chuẩn với trung bình là 158.0cm và độ lệch chuẩn 5.0cm:
 - Tính xác suất phụ nữ tham gia hoa hậu có chiều cao từ 168 – 190cm,
 - Tính khoảng tin cậy 95% (95%CI).
 - Tìm Outlier với khoảng ctin cậy 99%.
5. Bạn hiểu thế nào về giá trị P – values và ảnh hưởng của nó lên các nghiên cứu/đánh giá?
6. Công ty dược phẩm X cho ra mắt sản phẩm thuốc B và cho rằng nó có tác dụng điều trị đau dạ dày tốt hơn thuốc A hiện đang bán trên thị trường. Công ty X tiến hành nghiên cứu lâm sàng và cho kết quả như sau:

	Thuốc A	Thuốc B
Điều trị khỏi bệnh	367	409
Điều trị không khỏi	135	94

- Theo bạn, thiết kế nghiên cứu nào phù hợp cho nghiên cứu này?
 - Thuốc B liệu có tác dụng tốt hơn thuốc A?
 - Theo bạn, nếu công ty X bán sản phẩm ra thị trường thì công ty X cần bổ sung thông tin gì?
7. Để đánh giá chất lượng giảng dạy và cảm nhận của học viên khóa AIO 2021, 1 bảng câu hỏi sử dụng thang đo Likert (1. Rất không hài lòng, 2. Không hài lòng, 3. Trung dung, 4. Hài lòng, 5. Rất hài lòng) được dùng để các học viên đánh giá. Bạn sẽ phân tích kết quả này như thế nào để biết được đánh giá của học viên?
 8. **Câu hỏi bổ sung cho các bạn thích nghiên cứu khoa học: Theo mục tiêu nghiên cứu, các thiết kế nghiên cứu lâm sàng hiện nay thường được sử dụng là gì?**