

**Trường Chính sách công và Quản lý Fulbright**

Học kỳ Xuân năm 2020

**THỐNG KÊ VÀ PHÂN TÍCH DỮ LIỆU**

**BÀI TẬP 2**

**Ngày Phát: Thứ hai 09/03/2020**

**Ngày Nộp: Thứ tư 11/03/2020**

Học viên nộp bản điện tử bao gồm Word và Excel trước 13h20 trên ứng dụng Microsoft Teams

Học viên cần tính toán trên Excel và trình bày rõ ràng trên Word

**Câu 1:** Bài tập này vẫn sử dụng file dữ liệu NGTK.xlsx về các dữ liệu thống kê kinh tế xã hội của Việt Nam được lấy từ GSO (<https://www.gso.gov.vn/>) để trả lời các câu hỏi.

	Đơn vị	Diễn giải
<b>province</b>		Tỉnh
<b>region</b>		Vùng
<b>acreage</b>	km <sup>2</sup>	Diện tích
<b>population</b>	nghìn người	Dân số trung bình
<b>density</b>	người/km <sup>2</sup>	Mật độ dân số/km <sup>2</sup>
<b>labor</b>	nghìn người	Lực lượng lao động từ 15 tuổi trở lên
<b>scale</b>		Phân loại mật độ dân số (Thấp: mật độ < 400 người/km <sup>2</sup> Trung Bình: 400 < mật độ < 800 Cao: mật độ ≥ 800 người/km <sup>2</sup> )
<b>migrate</b>	%	Tỷ suất di cư thuần
<b>ind-index</b>	%	Chỉ số sản xuất công nghiệp
<b>s-market</b>	siêu thị	Số lượng siêu thị
<b>project</b>	dự án	Số dự án FDI
<b>capital</b>	triệu USD	Tổng vốn FDI đăng ký

Dựa vào bộ dữ liệu, anh/chị hãy tìm xác suất để:

- Chọn một tỉnh ngẫu nhiên thì tỉnh đó thuộc Trung du và miền núi phía Bắc?
- Chọn một tỉnh ngẫu nhiên thuộc Trung du và miền núi phía Bắc thì tỉnh đó mức phân loại mật độ dân số là Thấp?
- Chỉ được dựa vào kết quả của câu a và b, anh/chị hãy tính xác suất để chọn một tỉnh ngẫu nhiên thì tỉnh đó vừa thuộc Trung du và miền núi phía Bắc và vừa thuộc mức phân loại mật độ dân số là Thấp?
- Xác suất để chọn một tỉnh ngẫu nhiên có số dự án FDI đầu tư nằm trong khoảng từ 50 tới 300 dự án?
- Xác suất để chọn một tỉnh ngẫu nhiên mà số dự án FDI đầu tư nằm trong khoảng từ 50 tới 300 dự án hoặc vốn đầu tư nằm trong khoảng từ 1000 tới 5000 triệu USD và tỉnh này thuộc khu vực Đồng Bằng Sông Hồng là?

**Câu 2:**

VHLSS là bộ dữ liệu về tổng điều tra mức sống hộ gia đình của Việt Nam. Dựa vào bộ VHLSS tính được thu nhập bình quân của người dân Việt Nam vào năm 2018 khoảng 60 triệu VNĐ/năm và độ lệch chuẩn là 20 triệu VNĐ. Giả định thu nhập của người dân Việt Nam tuân theo phân phối chuẩn, các tỉ lệ trong mẫu VHLSS tương đương với tỉ lệ trong tổng thể. Biết rằng, dân số Việt Nam vào năm 2018 là 94.7 triệu người. Anh/chị hãy trả lời các câu hỏi sau cho năm 2018:

- Số người có thu nhập từ 80 triệu VNĐ/ năm trở lên?
- Số người có thu nhập từ 30 đến 80 triệu VNĐ/năm là?
- Có chuyên gia cho rằng, ít nhất 50% dân số có thu nhập bình quân hàng năm trong khoảng từ 50 tới 70 triệu VNĐ. Anh/chị nghĩ thế nào về nhận định trên?
- Hãy tính mức thu nhập mà ở đó 90% người dân có thu nhập thấp hơn mức này?

**Câu 3:**

Vấn sử dụng bộ dữ liệu VHLSS, với thu nhập bình quân đầu người của Việt Nam tính được là 60 triệu VNĐ/ năm. Học viên trả lời các câu hỏi sau:

- Một chuyên gia kinh tế nghi ngờ số liệu trên là không đúng vì cho rằng bộ dữ liệu VHLSS thu thập không đáng tin cậy nên đã khảo sát ngẫu nhiên một số người trên khắp Việt Nam thì kết quả về thu nhập bình quân hàng năm của những người được khảo sát thu được như sau:

65	56	65	57	62
62	58	60	65	55
59	57	65	56	56
61	60	58	58	58
64	64	63	56	61
56	62	63	60	59
63	60	65	62	61
63	58	63	56	55
59	64	60	63	60
60	58	65	61	63

Nếu kết quả lấy mẫu ngẫu nhiên là đáng tin cậy, anh/chị hãy phát biểu giả thuyết hợp lý và kiểm định xem với độ tin cậy 95% thì thu nhập bình quân của người Việt Nam được tính từ bộ dữ liệu VHLSS có chính xác không? (Anh/chị hãy giải bằng 2 cách là dùng trị kiểm định và p-value)

- Một chuyên gia kinh tế khác cho rằng thu nhập của người dân TP.HCM cao hơn thu nhập bình quân của Việt Nam. Nên đã khảo sát dữ liệu về thu nhập của 100 người dân TP.HCM bất kỳ thì có thu nhập trung bình là 63 triệu VNĐ/năm và độ lệch chuẩn 4 triệu VNĐ/năm. Hãy kiểm định nhận định của vị chuyên gia với độ tin cậy 95%?

**Câu 4:**

- Một doanh nghiệp bán lẻ nước ngoài muốn đầu tư vào Việt Nam nhưng đang phân vân nên đầu tư vào Hà Nội hay TP. Hồ Chí Minh. Doanh nghiệp này thuê một công ty khảo sát, tư vấn để tìm hiểu xem thu nhập bình quân đầu người Hà Nội hay TP. Hồ Chí Minh cao hơn. Sau khi khảo sát mẫu 100 người mỗi tỉnh thì thu được kết quả sau:

- Thu nhập bình quân đầu người của mẫu thuộc Hà Nội là 80 triệu VNĐ/năm với độ lệch chuẩn 20 triệu VNĐ/năm.

- Thu nhập bình quân đầu người của mẫu thuộc TP. Hồ Chí Minh là 90 triệu VNĐ/năm với độ lệch chuẩn là 25 triệu VNĐ/năm.

Giả định quyết định đầu tư của doanh nghiệp bán lẻ nước ngoài này chỉ phụ thuộc vào thu nhập bình quân đầu người mỗi tỉnh.

Nếu kết quả khảo sát là đáng tin cậy, thì doanh nghiệp này đã đủ cơ sở để ra quyết định chưa với độ tin cậy 95%?

b. Chính phủ đang cân nhắc một chính sách đưa ra có làm tăng thu hút dự án đầu tư FDI hay không, nên đã thử nghiệm 2 năm ở 5 tỉnh thành lớn, kết quả như sau:

	2018 chưa có chính sách (đơn vị: dự án)	2019 đã có chính sách (đơn vị: dự án)
TPHCM	8320	8714
Hà Nội	5413	5120
Cần Thơ	93	110
Đà Nẵng	671	602
Hải Phòng	749	821

Giả định các điều kiện kinh tế vĩ mô của các tỉnh không có sự thay đổi giữa năm 2018 và 2019, đồng thời sự thay đổi FDI chỉ do tác động của chính sách đem lại.

Từ kết quả trên, anh/chị có tin rằng chính sách có hiệu quả với độ tin cậy 95% không? (gợi ý: học viên cần phát biểu đầy đủ giả thuyết sau đó kiểm định sự khác biệt cặp)