Distributions

Thursday, November 23, 2023 9:51 AM

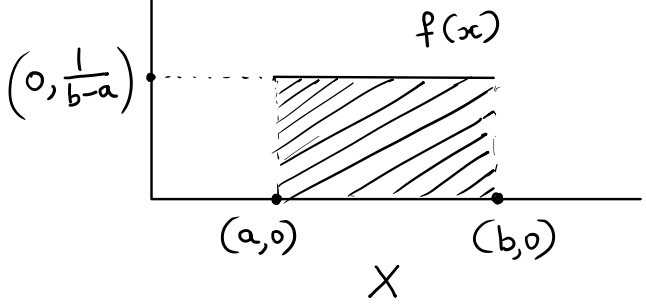
Uniform Distribution

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b, \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$$

$$\int_a^b f(x)\,dx = \int_a^b \frac{1}{b-a}\,dx$$

$$= \frac{1}{b-a}\left[x\right]_a^b = \frac{b-a}{b-a} = 1$$

$$A(\square) = b-a \cdot \frac{1}{b-a} = 1$$



$E(X)$ is Arithmetic mean of Prob. Distribution

$$E(X) = \sum x_i p_i \qquad E(X^2) = \sum x_i^2 p_i$$

$$E[g(X)] = \sum g(x_i) p_i$$

$$Var(X) = E(X^2) - [E(X)]^2$$

1. Consider a random variable with the following probability distribution:

X=r (or x_i)	P(X=r) or (p_i)
0	0.1
1	0.2
2	0.3
3	0.3
4	0.1

Find the following

a) P(X<=2)
b) P(1<X≤3)
c) P(X>0)
d) P(X>3 | X>2)
e) E(X) $= 2.1$
f) Var(X) $= \sum x_i^2 p_i - \left(\sum x_i p_i\right)^2 = 1.29$

$P(A \cup B) = P(A) + P(B)$
A & B
disjoint

a) $P(X \le 2) = P[X=0 \text{ or } X=1 \text{ or } X=2]$
$= P[X=0] + P[X=1] + P[X=2] = 0.6$

b) $P(1 < X \le 3) = P[X=2] + P[X=3] = 0.6$

c) $P[X>0] = 1 - P[X=0] = 0.9$

d) $P(X>3 \mid X>2) = \frac{P(X>3 \cap X>2)}{P[X>2]} = \frac{P(X>3)}{P[X>2]} = \frac{0.1}{0.4} = 0.25$

Joint Prob. Distribution

Aspiration Transmission	Naturally Aspirated	Turbocharged/Supercharged	All
Automatic	0.316434	0.347902	0.664336
CVT	0.063811	0.010490	0.074301
Manual	0.076923	0.184441	0.261364
All	0.457168	0.542832	1.000000

Marginal Prob Dist for Transmission

Marginal Prob. distribution for Aspiration

$P(\text{Auto Trans} \cap \text{Nat Asp}) = 0.316434$

$P(\text{Manual} \cap \text{Turbo}) = 0.184441$

$P(\text{Nat Asp.}) = 0.457168$ (Marginal) $P(\text{Turbo}) = 0.542832$

	A1	A2	A3	A4	Marginal
B1	P(A1∩B1)	P(A2∩B1)	P(A3∩B1)	P(A4∩B1)	P(B1)
B2	P(A1∩B2)	P(A2∩B2)	P(A3∩B2)	P(A4∩B2)	P(B2)
B3	P(A1∩B3)	P(A2∩B3)	P(A3∩B3)	P(A4∩B3)	P(B3)
Marginal	P(A1)	P(A2)	P(A3)	P(A4)	1

$\frac{a}{a+b+c}$, $\frac{b}{a+b+c}$, $\frac{c}{a+b+c}$ } sum = 1

$$P(B_1 \mid A_1) = \frac{P(A_1 \cap B_1)}{P(A_1)}$$

$$P(B_2 \mid A_1) = \frac{P(A_1 \cap B_2)}{P(A_1)}$$

$$P(B_3 \mid A_1) = \frac{P(A_1 \cap B_3)}{P(A_1)}$$

Conditional Prob. dist. for A_1

0.1	0.1/0.9
0.5	0.5/0.9
0.3	0.3/0.9
0.9	1

Aspiration Transmission	Naturally Aspirated	Turbocharged/Supercharged	All
Automatic	0.316434	0.347902	0.664336
CVT	0.063811	0.010490	0.074301
Manual	0.076923	0.184441	0.261364
All	0.457168	0.542832	1.000000

Conditional dist for Nat. Aspirated

$P(\text{Auto} \mid \text{NA}) = 0.316434 / 0.457168$ } sum = 1

$P(\text{Auto} \mid NA) = 0.316434 / 0.457168$

$P(\text{CVT} \mid NA) = 0.063811 / 0.457168$

$P(\text{Man} \mid NA) = 0.076923 / 0.457168$

} sum = 1

```
In [64]: pd.crosstab(index=survey['Exer'],
    ...:                columns=survey['Smoke'],
    ...:                margins=True, normalize=True)
Out[64]:
Smoke      Heavy     Never     Occas     Regul       All
Exer
Freq    0.032864  0.408451  0.056338  0.042254  0.539906
Some    0.014085  0.394366  0.018779  0.032864  0.460094
All     0.046948  0.802817  0.075117  0.075117  1.000000
```

$P(\text{Exer}_{\text{Freq}} \cap \text{Smoke}_{\text{Heavy}}) = 0.032864$

$P(\text{Exer}_{\text{Some}} \cap \text{Smoke}_{\text{Occas}}) = 0.018779$

$P(\text{Heavy Smoker}) = 0.046948$

```
In [65]: pd.crosstab(index=survey['Exer'],
    ...:                columns=survey['Smoke'],
    ...:                margins=True, normalize='columns')
Out[65]:
Smoke   Heavy     Never   Occas    Regul       All
Exer
Freq      0.7  0.508772    0.75  0.5625  0.539906
Some      0.3  0.491228    0.25  0.4375  0.460094
```

By Columns

$P(\text{Exer Freq} \mid \text{Heavy Smoke})$: There are 70% people among the heavy smokers who exercise frequently

$P(\text{Exer Some} \mid \text{Heavy Smoker})$: There are 30% people among the heavy smokers who exercise sometimes