Introduction to Mathematical Statistics I

Extra Credit

Freya Zou

1. Determine whether the following statements are true or false. Justify your answer.

(i) If A, B are independent, then A and B^c (the complement of B) are also independent.

True

If A, B are independent $P(A \cap B) = P(A)P(B)$

The complement of B is B^c, and we know: $P(A \cap B^c) = P(A) - P(A \cap B)$

we can substitute:

$$P(A \cap B^c) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c).$$

Thus we know $P(A \cap B^c) = P(A)P(B^c)$, showing A and B^c are independent

(ii) For a discrete random variable Y we have that $E(Y^2) \geq [E(Y)]^2$.

True:

According to variance property: $Var(Y) = E(Y^2) - [E(Y)]^2$

we know Variance is always non-negative ($Var(Y) \geq 0$).

thus we get $E(Y^2) - [E(Y)]^2 \geq 0$ which implies:

$$E(Y^2) \geq [E(Y)]^2.$$

(iii) If $Y \sim \text{Geometric}(p)$, then the random variable $Y^* = Y - 1$ has mean $\frac{1-p}{p}$.

True:

For $Y \sim \text{Geometric}(p)$, Y represents the number of trails required for the first success, and It's mean is $E(Y) = \frac{1}{p}$

the number of failures before the success. Subtracting 1 from Y decreases the mean by 1:

$$E(Y^*) = E(Y) - 1 = \frac{1}{p} - 1 = \frac{1-p}{p}$$

thus, the mean of Y^* is $\frac{1-p}{p}$.

2. A quality-control program at a plastic bottle production line involves inspecting finished bottles for flaws such as microscopic holes. The probability that a bottle has a flaw is 0.002. If a bottle has a flaw, the probability that it will fail the inspection is 0.995. If a bottle does not have a flaw, the probability that it will pass the inspection is 0.990.

   (i) If a bottle does not have a flaw, what is the probability that it will fail the inspection?

   (ii) What is the probability that a randomly selected bottle has a flaw and fails the inspection?

   (iii) If a bottle fails the inspection, what is the probability that it has a flaw?

Given: $P(\text{flaw}) = 0.002$, $P(\text{fail}|\text{flaw}) = 0.995$

$P(\text{pass}|\text{no flaw}) = 0.990$, $P(\text{no flaw}) = 1 - 0.002 = 0.998$

i) $P(\text{fail}|\text{no flaw}) = 1 - P(\text{Pass}|\text{no flaw}) = 1 - 0.990 = 0.01$

ii) $P(\text{flaw} \cap \text{fail}) = P(\text{fail}|\text{flaw}) \cdot P(\text{flow}) = 0.995 * 0.002 = 0.00199$

iii) $$P(\text{flaw}|\text{fail}) = \frac{P(\text{fail}|\text{flaw})\,P(\text{flaw})}{P(\text{fail}|\text{flaw})\,P(\text{flaw}) + P(\text{fail}|\text{no flaw})\,P(\text{no flaw})}$$

$$= \frac{0.995 * 0.002}{0.995 * 0.002 + 0.01 * 0.998}$$

$$= \frac{0.00199}{0.00199 + 0.00998} = \frac{0.00199}{0.01197}$$

$$= 0.1662489$$

3. Consider an experiment consisting in rolling a fair die twice. We can represent the possible outcomes by ordered pairs. For instance $(1, 3)$ means that we obtained a 1 in the first roll and a 3 in the second roll. Since $(1, 3)$ and $(3, 1)$ are different outcomes we have total of 36 (equally likely) possible outcomes for this experiment. Define the random variable Y to be the sum of the numbers observed in the two rolls.

(i) Write down all the possible values that the random variable Y can take.

(ii) Find $P(Y = 6)$.

(iii) Suppose that somebody tells you that he performed the experiment and observed the value $Y = 4$. What is the probability that he obtained a 3 in the first roll?

i) Possible values that the random variable Y can take: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

ii) two rolls have sample space: $6 \cdot 6 = 36$

There are 5 outcomes where $Y=6$: $(3,3)$ $(2,4)$ $(4,2)$ $(1,5)$ $(5,1)$

thus: $P(Y=6) = \dfrac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}} = \dfrac{5}{36} \approx 0.13889$

iii) when $Y=4$: there are 3 situations are:

$(2,2)$ $(1,3)$ $(3,1)$

The First $Y=3$ has $(1,2)$

3) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = P(\text{First roll}=3 \mid Y=4) = \dfrac{P(\text{First roll}=3) \cap P(Y=4)}{P(Y=4)}$

when $Y=4$ have 3 outcomes: $(1,3)$, $(2,2)$, $(3,1)$ and we have sample space $6 \cdot 6 = 36$ thus $P(Y=4) = \dfrac{3}{36}$

we also can find $P(\text{first roll}=3) \cap P(Y=4)$, which means first roll is 3 and the sum is 4, This happens only one outcome $(3,1)$, thus $P(\text{first roll}=3) \cap P(Y=4) = 1/36$

$P(\text{first roll}=3 \mid Y=4) = \dfrac{P(\text{first roll}=3) \cap P(Y=4)}{P(Y=4)} = \dfrac{1/36}{3/36} = \dfrac{1}{3} \approx 0.333$

4. Let Y be a random variable with pdf

$$f(y) = \begin{cases} 0.2 & , \ -1 < y < 0 \\ 0.2 + cy & , \ 0 \le y \le 1 \\ 0 & , \ \text{elsewhere} \end{cases}$$

(i) Find the cdf $F(y)$ (Note: First you need to determine the value of c)

(ii) Find $P(Y < 1/2 | Y > -1/2)$.

(iii) Find the mean and variance of Y.

i) PDF $= \int_{-\infty}^{\infty} f(y)\,dy = 1 = \int_{-1}^{1} f(y)\,dy = \int_{-1}^{0} 0.2\,dy + \int_{0}^{1} (0.2 + cy)\,dy = 1$

$$\Rightarrow = \int_{-1}^{0} 0.2\,dy + \int_{0}^{1} 0.2\,dy + \int_{0}^{1} cy\,dy$$

$$\Rightarrow = 0.2 + 0.2 + \frac{c}{2} = 1$$

$$\Rightarrow \boxed{C = 1.2}$$

now, the PDF: $f(y) = \begin{cases} 0.2 & , \ -1 < y < 0 \\ 0.2 + 1.2y & , \ 0 \le y \le 1 \\ 0 & , \ \text{elsewhere} \end{cases}$

1. For $y < -1$: $F(y) = 0$, since $f(y) = 0$ outside $[-1, 1]$.

2. For $-1 \le y < 0$: $F(y) = \int_{-1}^{y} 0.2\,dt = 0.2(y+1) = 0.2y + 0.2$

3. For $0 \le y \le 1$: $F(y) = F(0) + \int_{0}^{y} (0.2 + 1.2dt)\,dt$

$$= 0 + \int_{0}^{y} 0.2\,dt + \int_{0}^{y} 1.2t\,dt$$

$$= 0.2y + 0.6y^2$$

Therefore, $F(y) = 0.2 + 0.2y + 0.6y^2 = 0.2 + 0.2y + 0.6y^2$

4. For $y > 1$: $F(y) = 1$. since probability is 1.

thus. CDF for Fy is: $\boxed{F(y) = \begin{cases} 0 & y < -1 \\ 0.2y + 0.2 & -1 \le y < 0 \\ 0.2 + 0.2y + 0.6y^2 & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}}$

ii) $P(Y<\frac{1}{2} \mid Y>-\frac{1}{2}) = \dfrac{P(-\frac{1}{2}<Y<\frac{1}{2})}{P(Y>-\frac{1}{2})}$

$$= \frac{F(\frac{1}{2}) - F(-\frac{1}{2})}{1-F(-\frac{1}{2})}$$

no we can subsitute, $y=-1/2$, using $F(y)=0.2y+0.2 \, (-1 \le y < 0)$:

$F(-\frac{1}{2}) = 0.2(-\frac{1}{2}) + 0.2 = 0.1 + 0.2 = 0.1$

and when $y=\frac{1}{2}$ we can subsitute in $F(y) = 0.2+0.2y+0.6y^2 \, (0 \le y \le 1)$

then we can get: $F(\frac{1}{2}) = 0.2+0.2(\frac{1}{2})+0.6(\frac{1}{2})^2 = 0.2+0.1+0.6(0.25)$
$= 0.45$

we also need to find $1-F(-\frac{1}{2}) = 1-0.1 = 0.9$ thus we know:

$$P(Y<\frac{1}{2} \mid Y>-\frac{1}{2}) = \frac{F(\frac{1}{2})-F(-\frac{1}{2})}{1-F(-\frac{1}{2})} = \frac{0.45-0.1}{0.9} \approx 0.39$$

iii)

$$\mu = E(Y) = \int_{-1}^{1} y f(y)\,dy = \int_{-1}^{0} y(0.2)\,dy + \int_{0}^{1} y(0.2+1.2y)\,dy$$

$$= 0.2\int_{-1}^{0} y\,dy + \int_{0}^{1} 0.2y\,dy + \int_{0}^{1} 1.2y^2\,dy$$

$$= -0.1 + 0.1 + 0.4$$

$$= 0.4$$

$$E(Y^2) = \int_{-1}^{0} y^2(0.2)\,dy + \int_{0}^{1} y^2(0.2+1.2y)\,dy$$

$$= 0.06666 + 0.36666 = 0.4333$$

$$var(Y) = E(Y^2) - [E(Y)]^2 = 0.4333 - (0.4)^2 \approx 0.27333$$

thus we get mean is 0.4 and variance is 0.27333

5. Suppose that $Y \sim \text{Gamma}(\alpha, \beta)$. That is, Y has pdf

$$f(y) = \begin{cases} \frac{y^{\alpha-1}e^{-y/\beta}}{\Gamma(\alpha)\beta^{\alpha}} & , 0 < y < \infty \\ 0 & , \text{otherwise} \end{cases}$$

where $\alpha, \beta > 0$.

(i) Verify that $\int_{\infty}^{-\infty} f(y)dy = 1$. (Hint: Recall the definition of a gamma function given in class)

(ii) Without using moment generating functions, find $E(Y)$ and $\text{Var}(Y)$.

(iii) Find the moment generating function of Y and use it to verify your results in part (ii).

i) since $f(y)$ is defined as zero for $y \leq 0$, we only need to verify the intergral over $(0, \infty)$:

$$\int_0^\infty f(y)dy = \int_0^\infty \frac{y^{\alpha-1}e^{-y/\beta}}{\Gamma(\alpha)\beta^{\alpha}} dy$$

$$= \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^\infty y^{\alpha-1} e^{-y/\beta} dy$$

Now we can Subsitude $u = \frac{y}{\beta}$, $y = \beta u$ and $dy = \beta du$:

$$\int_0^\infty y^{\alpha-1} e^{-y/\beta} dy = \int_0^\infty (\beta u)^{\alpha-1} e^{-u} \beta du$$

$$= \beta^{\alpha} \int_0^\infty u^{\alpha-1} e^{-u} du$$

Now we Subsitute back into originaly the original expression:

$$\int_0^\infty f(y)dy = \frac{1}{\Gamma(\alpha)\cancel{\beta^{\alpha}}} \cdot \cancel{\beta^{\alpha}} \int_0^\infty u^{\alpha-1} e^{-u} du.$$

$$= \frac{1}{\Gamma(\alpha)} \boxed{\int_0^\infty u^{\alpha-1} e^{-u} du} = \Gamma(\alpha)$$

by the definition of the Gamma distribution

$$= \frac{\Gamma(\alpha)}{\Gamma(\alpha)} = 1$$

Therefore: $\int_0^\infty f(y)dy = \frac{\Gamma(\alpha)}{\Gamma(\alpha)} = 1$, this veriey that $f(y)$ is a valid probability density function

ii) • Mean $E(Y)=\alpha\beta$:

The expected value of a random variable Y is given by:

$E(Y)=\int_0^\infty y f(y)\,dy$ where $f(y)$ is Gamma distribution:

$f(y)=\dfrac{y^{\alpha-1}e^{-y/\beta}}{\Gamma(\alpha)\beta^{\alpha}}$, subsituting the PDF into the formula for $E(Y)$:

$$E(Y)=\int_0^\infty y\,\frac{y^{\alpha-1}e^{-y/\beta}}{\Gamma(\alpha)\beta^{\alpha}}\,dy=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\int_0^\infty y^{\alpha}e^{-y/\beta}\,dy$$

Change of variable: let $u=\frac{y}{\beta}$ so $y=\beta u$ and $dy=\beta du$:

$$E(Y)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\int_0^\infty(\beta u)^{\alpha}e^{-u}\,du$$

$$=\frac{\beta^{\alpha+1}}{\Gamma(\alpha)\beta^{\alpha}}\int_0^\infty u^{a}e^{-u}\,du$$

$$=\frac{\beta}{\Gamma(\alpha)}\int_0^\infty u^{a}e^{-u}\,du$$

Since the intergoal of $u^{a}e^{-u}$ over u from $[0,\infty]$ equals $\Gamma(\alpha+1)$:

$$E(Y)=\frac{\beta\Gamma(\alpha+1)}{\Gamma(\alpha)}=\alpha\beta$$

• Variance $Var(Y)$: $Var(Y)=E(Y^2)-[E(Y)]^2$

$$E(Y^2)=\int_0^\infty y^2 f(y)\,dy=\int_0^\infty y^2\,\frac{y^{\alpha-1}e^{-y/\beta}}{\Gamma(\alpha)\beta^{\alpha}}\,dy$$

$$=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\int_0^\infty y^{\alpha+1}e^{-y/\beta}\,dy$$

Change of variable, $u=y/\beta$, $y=\beta u$, $dy=\beta du$:

$$E(Y^2)=\frac{\beta^{\alpha+2}}{\Gamma(\alpha)}\int_0^\infty(\beta u)^{\alpha+1}e^{-u}\beta\,du=\frac{\beta^{\alpha+2}\Gamma(\alpha+2)}{\Gamma(\alpha)}$$

$\Rightarrow \alpha(\alpha+1)\beta^2$ (using the property $\Gamma(\alpha+2)=(\alpha+1)\alpha\Gamma(\alpha)$)

$$Var(Y)=\alpha(\alpha+1)\beta^2-(\alpha\beta)^2=\alpha\beta^2$$

thus, $Var(Y)=\alpha\beta^2$

iii) The moment generating function (MGF) of a Gamma random variable $Y \sim Gamma(\alpha, \beta)$ is Given by $M_Y(t) = E(e^{tY})$

then MGF is: $\boxed{M_Y(t) = (1-\beta t)^{-\alpha}, \text{ for } t < \frac{1}{\beta}}$

verifying $E(Y)$ and $Var(Y)$ using the MGF:

1. Mean $E(Y)$:

The mean $E(Y)$ can be found by differentiating the MGF with respect to t and evaluating at $t=0$:

$$E(Y) = M'_Y(0)$$

differentiate $M_Y(t) = (1-\beta t)^{-\alpha}$ with respect to t:

$$M'_Y(t) = \alpha\beta(1-\beta t)^{-\alpha-1}$$

Subsitute $t=0$:

$$M'_Y(t) = \alpha\beta$$

so, $E(Y) = \alpha\beta$, which agrees with the result from part (ii)

2. variance $Var(Y)$:

To find $Var(Y)$, we need $E(Y^2) = M''_Y(0)$ and then use

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

differentiate $M_Y(t) = (1-\beta t)^{-\alpha}$ with respect to t:

$M''_Y(t) = \alpha\beta(\alpha+1)\beta(1-\beta t)^{-\alpha-2} = \alpha(\alpha+1)\beta^2(1-\beta t)^{-\alpha-2}$

Subsitute $t=0$: $M''_Y(0) = \alpha(\alpha+1)\beta^2$, thus $E(Y^2) = \alpha(\alpha+1)\beta^2$

Now. $Var(Y) = E(Y^2) - [E(Y)]^2 = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2$

$= \alpha\beta^2$

This Confirms that $Var(Y) = \alpha\beta^2$, which matches the result in part (ii)