STA 561 Homework 1 Solutions

Problem 1

1.1

Answer:

Denote knowing C/C++ as event C, Fortran as event F.

(a) $P(\bar{F}) = 1 - P(F) = 0.4 = 40\%$

(b) $P(\bar{F} \cap \bar{C}) = 1 - P(F \cup C) = 1 - [P(F) + P(C) - P(F \cap C)] = 1 - 0.6 - 0.7 + 0.5 = 0.2 = 20\%$

(c) $P(C \cap \bar{F}) = P(C) - P(C \cap F) = 0.7 - 0.5 = 0.2 = 20\%$

(d) $P(F \cap \bar{C}) = P(F) - P(F \cap C) = 0.6 - 0.5 = 0.1 = 10\%$

(e) $P(C|F) = P(C \cap F)/P(F) = 0.5/0.6 = 83.33\%$

(f) $P(F|C) = P(C \cap F)/P(C) = 0.5/0.7 = 71.43\%$

1.2

Answer:

Denote the discovery of error by the 5 independent tests as event A, B, C, D, and E, respectively.

(a) $P\{\text{discovered by at least one test}\} = P(A \cup B \cup C \cup D \cup E) = 1 - P(\bar{A} \cap \bar{B} \cap \bar{C} \cap \bar{D} \cap \bar{E}) = 1 - P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C}) \cdot P(\bar{D}) \cdot P(\bar{E}) = 1 - 0.9 \times 0.8 \times 0.7 \times 0.6 \times 0.5 = 0.8488 = 84.88\%$

(b) $P\{\text{discovered by at least two tests}\} = P\{\text{discovered by at least one test}\} - P\{\text{discovered by one test}\} = 0.8488 - 0.1 \times 0.8 \times 0.7 \times 0.6 \times 0.5 - 0.9 \times 0.2 \times 0.7 \times 0.6 \times 0.5 - 0.9 \times 0.8 \times 0.3 \times 0.6 \times 0.5 - 0.9 \times 0.8 \times 0.7 \times 0.4 \times 0.5 - 0.9 \times 0.8 \times 0.7 \times 0.6 \times 0.5 = 0.4774 = 47.74\%$

(c) $P\{\text{discovered by all five tests}\} = P(A \cap B \cap C \cap D \cap E) = P(A) \cdot P(B) \cdot P(C) \cdot P(D) \cdot P(E) = 0.1 \times 0.2 \times 0.3 \times 0.4 \times 0.5 = 0.0012 = 0.12\%$

Problem 2

Show that $a^T(rb + c) = r(a^Tb) + a^Tc$ for random variables a, b, c and scalar r.

Suppose $a^T = [1, 2, 3]$ and $b^T = [2, 3, 4]$ and $c^T = [4, 5, 6]$ and $r = 2$, then we have:

$$\begin{aligned}
a^T(rb + c) &= a_1(rb_1 + c_1) + a_2(rb_2 + c_2) + a_3(rb_3 + c_3) \\
&= \sum_{i=1}^{3} a_i(rb_i + c_i) \\
&= 1 \times (2 \times 2 + 4) + 2 \times (2 \times 3 + 5) + 3 \times (2 \times 4 + 6) \\
&= 1 \times 8 + 2 \times 11 + 3 \times 14 \qquad\qquad = 72
\end{aligned}$$

We also have

$$\begin{aligned}
ra^Tb &= r(a_1b_1 + a_2b_2 + a_3b_3) \\
&= 2\sum_{i=1}^{3} a_ib_i \\
&= 2(1 \times 2 + 2 \times 3 + 3 \times 4) \\
&= 2(2 + 6 + 12) \\
&= 2(20) \\
&= 40
\end{aligned}$$

anbd we previously calculated $a^Tc = 32$. Thus:

$$\begin{aligned}
a^T(rb + c) &= 72 \\
&= r(a^Tb) + a^Tc
\end{aligned}$$

Problem 3

3.1

Show that $E(cx) = cE(x)$ for constant c and random variable x.

Suppose x is a random variable with probability mass function:

$$p(x) = \begin{cases} 1/2 & \text{if } x = 1 \\ 1/2 & \text{if } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

and $c = 2$, meaning c has probability mass function:

$$p(c) = \begin{cases} 1 & \text{if } c = 2 \\ 0 & \text{otherwise} \end{cases}$$

Consider the event $x = 1, c = 2$. This always happens for c, and this happens with probability $1/2$ for x, so the probability of the event is $1/2 = 1 \times 1/2$. As a result, $p(c, x)$ has the following joint distribution:

$$p(c,x) = \begin{cases} 1/2 & \text{if } x = 1, c = 2 \\ 1/2 & \text{if } x = 2, c = 2 \\ 0 & \text{if } x \neq 1 \text{ or } 2, \text{ or } c \neq 2 \end{cases}$$

It follows that any random variable is statistically independent of a constant random variable. Then:

$$\begin{aligned} E(cx) &= \sum_{x \in \{1,2\}} \sum_{c \in \{2, \neq 2\}} cxp(c,x) \\ &= \sum_{x \in \{1,2\}} 2 \times xp(c = 2, x) + \sum_{x \in \{1,2\}} c \times xp(c \neq 2, x) \\ &= \sum_{x \in \{1,2\}} 2 \times xp(c = 2, x) + \sum_{x \in \{1,2\}} c \times x \times 0 \\ &= 2 \times 1p(c = 2, x = 1) + 2 \times 2p(c = 2, x = 2) \\ &= 2 \times 1/2 + 4 \times 1/2 \\ &= 3 \end{aligned}$$

Next,

$$\begin{aligned} cE(x) &= 2 \sum_{x \in \{1,2\}} xp(x) \\ &= 2 \times (1 \times 1/2 + 2 \times 1/2) \\ &= 2 \times (1/2 + 1) \\ &= 3 \end{aligned}$$

Thus:

$$E(cx) = cE(x)$$

3.1 (b)

Show that $Var(a) = 0$ for constant a.

Using the definition of variance and the fact that $E(a) = a$ for constant random variable a, we have:

$$\begin{aligned} Var(a) &= E\Big[(a - E(a))(a - E(a))\Big] \\ &= E\Big[(a - a)(a - a)\Big] \\ &= E(0 \times 0) \\ &= 0 \end{aligned}$$

3.1(c)

Show that $Cov(x, y) = E(xy)$ if $E(x) = E(y) = 0$ for random variables x, y.

Using the definition of covariance:

$$\begin{aligned} Cov(x, y) &= E\Big[(x - E(x))(y - E(y))\Big] \\ &= E\Big[(x - 0)(y - 0)\Big] \\ &= E(xy) \end{aligned}$$

3.2

Suppose we have a discrete random variable x that takes on values $1, 3$ with probability $(1/3, 2/3)$, and, independently, z that takes on values $5, 10$ with probability $(1/2, 1/2)$. What is expected value $E[x-z]$ and $Var(x-z)$?

Using the definition of expected value, we have:

$$\begin{aligned}
E(x-z) &= \sum_{x\in\{1,3\}} \sum_{z\in\{5,10\}} (x-z)p(x)p(z) \\
&= \sum_{x\in\{1,3\}} \sum_{z\in\{5,10\}} xp(x)p(z) - \sum_{x\in\{1,3\}} \sum_{z\in\{5,10\}} zp(x)p(z) \\
&= \sum_{x\in\{1,3\}} xp(x)p(z=5) + \sum_{x\in\{1,3\}} xp(x)p(z=10) \\
&- \sum_{z\in\{5,10\}} zp(x=1)p(z) - \sum_{z\in\{5,10\}} zp(x=3)p(z) \\
&= 1 \times 1/3 \times 1/2 + 3 \times 2/3 \times 1/2 + 1 \times 1/3 \times 1/2 + 3 \times 2/3 \times 1/2 \\
&- 5 \times 1/3 \times 1/2 - 10 \times 1/3 \times 1/2 - 5 \times 2/3 \times 1/2 - 10 \times 2/3 \times 1/2 \\
&= 1/6 + 6/6 + 1/6 + 6/6 \\
&- 5/6 - 10/6 - 10/6 - 20/6 \\
&= -31/6
\end{aligned}$$

Note that this is the same as $E(x) - E(z)$. Before finding the variances, let's call $r = x - E(x)$ and $s = z - E(z)$, then:

$$\begin{aligned}
Var(x-z) &= E\Big[(x - z - E(x-z))(x - z - E(x-z))\Big] \\
&= E\Big[(x - E(x) - (z - E(z)))(x - E(x) - (z - E(z)))\Big] \\
&= E\Big[(r-s)(r-s)\Big] \\
&= E(r^2) + E(s^2) - 2E(rs) \\
&= E\Big[(x - E(x))(x - E(x))\Big] + E\Big[(z - E(z))(z - E(z))\Big] - 2E(rs) \\
&= Var(x) + Var(z) - 2E(rs)
\end{aligned}$$

Examining $E(rs)$:

$$\begin{aligned}
E(rs) &= E\Big[(x - E(x))(z - E(z))\Big] \\
&= Cov(x, z)
\end{aligned}$$

Because x and z are statistically independent, then $Cov(x, z) = 0$, giving:

$$Var(x-z) = Var(x) + Var(z)$$

Show that $Var(x) = E(x^2) - E(x)^2$ for any random variable x.

Using the definition of variance:

$$
\begin{aligned}
Var(x) &= E\Big[(x - E(x))(x - E(x))\Big] \\
&= E\Big[x^2 - xE(x) - E(x)x + E(x)^2\Big] \\
&= E(x^2) - E(xE(x)) - E(E(x)x) + E(E(x)^2) \\
&= E(x^2) - E(x)^2 - E(x)^2 + E(x)^2 \\
&= E(x^2) - E(x)^2
\end{aligned}
$$

Show that $Var(x) = E(x^2)$ if $E(x) = 0$ for any random variable x.

Using our previous result:

$$
\begin{aligned}
Var(x) &= E(x^2) - E(x)^2 \\
&= E(x^2) - 0^2 \\
&= E(x^2)
\end{aligned}
$$

3.3 (a)

Find the expectation of $\bar{x} = \sum_{i=1}^{n} x_i/n$ for any set of random variables $x_1, x_2, \ldots x_n$, not necessarily independent and identically distributed (IID). What if they are IID?

Suppose we have a set of random variables $x_1, x_2, \ldots x_n$. Then the expectation of $\bar{x}$ is:

$$
\begin{aligned}
E(\bar{x}) &= E(\sum_{i=1}^{n} x_i/n) \\
&= \sum_{i=1}^{n} E(x_i/n) \\
&= \sum_{i=1}^{n} E(x_i)/n \\
&= \frac{E(x_1) + E(x_2) + \cdots + E(x_n)}{n}
\end{aligned}
$$

If they are IID then they have the same expectation, let's call it μ, giving:

$$\begin{aligned}\frac{E(x_1) + E(x_2) + \cdots + E(x_n)}{n} &= \frac{\mu + \mu + \ldots \mu}{n} \\ &= \frac{n\mu}{n} \\ &= \mu\end{aligned}$$

3.3 (b)

Find the variance of $\bar{x} = \sum_{i=1}^{n} x_i/n$ for any set of random variables $x_1, x_2, \ldots x_n$, not necessarily independent and identically distributed (IID). What if they are IID?

Suppose we have a set of random variables $x_1, x_2, \ldots x_n$. Let's choose three random variables with $E(x_1) = \mu_1$ and $Var(x_1) = \sigma_1^2$, $E(x_2) = \mu_2$ and $Var(x_2) = \sigma_2^2$, and $E(x_3) = \mu_3$ and $Var(x_3) = \sigma_3^2$. Let's also use $r_1 = x_1 - E(x_1)$, $r_2 = x_2 - E(x_2)$, and $r_3 = x_3 - E(x_3)$. Then we have:

$$
\begin{aligned}
Var(\bar{x}) &= Var(\sum_{i=1}^{3} x_i/3) \\
&= E\Big[\Big(\frac{x_1+x_2+x_3}{3} - E(\frac{x_1+x_2+x_3}{3})\Big)\Big(\frac{x_1+x_2+x_3}{3} - E(\frac{x_1+x_2+x_3}{3})\Big)\Big] \\
&= E\Big[\Big(\frac{x_1 - E(x_1)}{3} + \frac{x_2 - E(x_2)}{3} + \frac{x_3 - E(x_3)}{3}\Big)^2\Big] \\
&= E\Big[\Big(\frac{r_1}{3} + \frac{r_2}{3} + \frac{r_3}{3}\Big)^2\Big] \\
&= E\Big[\Big(\frac{1}{3}(r_1 + r_2 + r_3)\Big)^2\Big] \\
&= \frac{1}{9}E\Big[\Big(r_1 + r_2 + r_3\Big)^2\Big] \\
&= \frac{1}{9}E\Big[r_1^2 + r_2^2 + r_3^2 + 2r_1r_2 + 2r_1r_3 + 2r_2r_3\Big] \\
&= \frac{1}{9}\Big[E(r_1^2) + E(r_2^2) + E(r_3^2) + E(2r_1r_2) + E(2r_1r_3) + E(2r_2r_3)\Big] \\
&= \frac{1}{9}\Big[Var(x_1) + Var(x_2) + Var(x_3) + 2Cov(x_1, x_2) + 2Cov(x_1, x_3) + 2Cov(x_2, x_3)\Big] \\
&= \frac{1}{9}\Big[\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + 2Cov(x_1, x_2) + 2Cov(x_1, x_3) + 2Cov(x_2, x_3)\Big]
\end{aligned}
$$

We can generalize this for n random variables

$$Var(\bar{x}) = \frac{1}{n^2}\Big[\sum_{i=1}^{n} \sigma_i^2 + 2\sum_{i \neq j} Cov(x_i, x_j)\Big]$$

If they are identically distributed then $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$ and $Cov(x_1, x_2) = Cov(x_1, x_3) = Cov(x_2, x_3)$, giving:

$$
\begin{aligned}
Var(\bar{x}) &= \frac{1}{9}\Big[\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + 2Cov(x_1, x_2) + 2Cov(x_1, x_3) + 2Cov(x_2, x_3)\Big] \\
&= \frac{1}{9}\Big[3\sigma_1^2 + 6Cov(x_1, x_2)\Big]
\end{aligned}
$$

Again, generalizing gives:

$$Var(\bar{x}) = \frac{1}{n^2}\Big[n\sigma^2 + 2nCov(x_i, x_j)\Big]$$
$$= \frac{1}{n}\Big[\sigma^2 + 2Cov(x_i, x_j)\Big]$$

Finally, if we assume both independence and identical distribution, the covariances go away, giving:

$$Var(\bar{x}) = \frac{\sigma^2}{n}$$

This is the typical variance of the sample mean that you learn, or the standard error, which is the square root of this:

$$SE(\bar{x}) = \sqrt{Var(\bar{x})} = \frac{\sigma}{\sqrt{n}}$$

Problem 4

Reference solutions can be found here (for python):

https://jakevdp.github.io/PythonDataScienceHandbook/05.06-linear-regression.html

https://github.com/justmarkham/DAT4/blob/master/notebooks/08_linear_regression.ipynb

Your answers may vary.