**Statistics 394, Problem Set 8 Solutions**

Wellner; 3/1/2000

1. K, 4.2, # 3, page 256. If $E(X) = 4$ and $E(X^2) = 21$, find:
(a) $E(X^2 - 3X + 2)$; (b) $E(X+1)^2$; (c) $E(X - EX)^2$; (d) $E(X^2) - (EX)^2$.

   **Solution:** (a) By Fact 1,

$$\begin{aligned} E(X^2 - 3X + 2) &= E(X^2) - 3E(X) + 2 \\ &= 21 - 3 \cdot 4 + 2 = 11\,. \end{aligned}$$

   (b) By algebra and Fact 1,

$$E(X+1)^2 = E(X^2 + 2X + 1) = E(X^2) + 2E(X) + 1 = 21 + 8 + 1 = 30\,.$$

   (c) and (d) $E(X - EX)^2 = E(X^2) - (EX)^2 = 21 - 4^2 = 5$.

2. K, 4.2, # 4, page 256. Suppose that $E(X) = \mu$. Show that:
(a) $E(X-\mu)^2 = E(X^2) - \mu^2$.
(b) For any $a$, $E(X-a)^2 = E(X-\mu)^2 + (\mu - a)^2$.
(c) $E(X-a)^2$ is minimized by $a = \mu$.

   **Solution:** (a) This is just our computational formula for $Var(X)$: by algebra and Fact 1,

$$\begin{aligned} E(X-\mu)^2 &= E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2\,. \end{aligned}$$

   (b) By writing

$$(X-a)^2 = (X - \mu + \mu - a)^2 = (X-\mu)^2 + 2(X-\mu)(\mu - a) + (\mu - a)^2\,,$$

   and then using Fact 1, it follows that

$$\begin{aligned} E(X-a)^2 &= E(X-\mu)^2 + 2(\mu - a)E(X-\mu) + (\mu - a)^2 \\ &= E(X-\mu)^2 + 2(\mu - a)\cdot 0 + (\mu - a)^2 \\ &= E(X-\mu)^2 + (\mu - a)^2\,. \end{aligned}$$

   (c) Note that from (b) we have

$$E(X-a)^2 \geq E(X-\mu)^2$$

   with equality if and only if $a = \mu$. Thus $E(X-a)^2$ is minimized by $a = \mu$.

3. K, 4.2, # 8, page 257. In Example 4.1.11 we found that if $X$ has the Poisson distribution with parameter $\lambda$, then $E(X) = \lambda$.
(a) Show that $E(X(X-1)) = \lambda^2$.
(b) From part (a) and $E(X)$, find $E(X^2)$.
(c) Use the result of Exercise 4(a) to find $Var(X) = E(X-\mu)^2$.

**Solution:** I did this problem in class earlier this quarter; but here it is again.
(a) Now

$$\begin{aligned} E[X(X-1)] &= \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=2}^{\infty} \frac{k(k-1)}{k!}\lambda^k e^{-\lambda} \\ &= \sum_{k=2}^{\infty} \lambda^2 \frac{\lambda^{k-2}}{(k-2)!}e^{-\lambda} = \lambda^2 \sum_{j=0}^{\infty} \frac{\lambda^j}{j!}e^{-\lambda} \\ &= \lambda^2 \cdot 1 = \lambda^2 . \end{aligned}$$

(b) Now $\lambda^2 = E[X(X-1)] = E(X^2 - X) = E(X^2) - E(X)$, so $E(X^2) = \lambda^2 + E(X) = \lambda^2 + \lambda$.
(c) It follows from (b) and our computational formula for $Var(X)$ that

$$Var(X) = E(X^2) - (EX)^2 = \lambda + \lambda^2 - \lambda^2 = \lambda .$$

4. K, 4.2, # 16, page 258. Let $X$ and $Y$ be independent random variables. Suppose that $E(X) = 2$, $E(X^2) = 6$, $E(Y) = 3$, and $E(Y^2) = 13$. Find the following:
(a) $E(X+Y)$. (b) $E(2XY)$. (c) $E(3X-Y)^2$. (d) $[E(3X-Y)]^2$.

**Solution:** (a) By Fact 2, $E(X+Y) = E(X) + E(Y) = 2 + 3 = 5$.
(b) By Fact 1 and independence of $X$ and $Y$,

$$E(2XY) = 2E(XY) = 2E(X) \cdot E(Y) = 2 \cdot 2 \cdot 3 = 12 .$$

(c) By algebra, Fact 2, and independence,

$$\begin{aligned} E(3X-Y)^2 &= E[9X^2 - 6XY + Y^2] = 9E(X^2) - 6E(XY) + E(Y^2) \\ &= 9 \cdot 6 - 6E(X) \cdot E(Y) + 13 \\ &= 54 - 6 \cdot 2 \cdot 3 + 13 = 31 . \end{aligned}$$

(d) $[E(3X-Y)]^2 = [3E(X) - E(Y)]^2 = [3 \cdot 2 - 3]^2 = 3^2 = 9$.

5. K, 4.3, # 1, page 268. Find $E(X)$ and $Var(X)$ if $X \sim$ Uniform$(a, b)$.

**Solution:** We did this one in class, but here it is again:
Solution 1.

$$E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^2}{2}\Big|_a^b = \frac{1}{2(b-a)}(b^2 - a^2) = (b+a)(b-a)/(2(b-a)) = (b+a)/2 ,$$

while

$$
\begin{aligned}
E(X^2) &= \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^3}{3} |_a^b = \frac{1}{3(b-a)} (b^3 - a^3) \\
&= (b^2 + ab + a^2)(b-a)/(3(b-a)) = (b^2 + ab + a^2)/3
\end{aligned}
$$

so that

$$
Var(X) = E(X^2) - (EX)^2 = (b^2+ab+a^2)/3 - (b+a)^2/4 = (b^2 - 2ab + a^2)/12 = (b-a)^2/12 \, .
$$

Solution 2. Note that if $U \sim \text{Uniform}(0,1)$, then $X \equiv (b-a)U + a \sim \text{Uniform}(a,b)$. Thus by Fact 1

$$
E(X) = (b-a)E(U) + a = (b-a)/2 + a = (b+a)/2 \, ,
$$

and

$$
Var(X) = (b-a)^2 Var(U) = (b-a)^2/12
$$

since

$$
E(U) = \int_0^1 u du = 1/2 \quad \text{and} \quad E(U^2) = \int_0^1 u^2 du = 1/3
$$

so that $Var(U) = 1/3 - (1/2)^2 = 1/12$.

6. K, 4.3, # 2, page 268. Let $X$ be a discrete random variable with possible values $1, 2, \ldots, n$ all equally likely. Find $E(X)$ and $Var(X)$.

**Solution:** Now $P(X = k) = 1/n$ for $k = 1, \ldots, n$, and hence

$$
E(X) = \sum_{k=1}^n k \frac{1}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2} \, ,
$$

$$
E(X^2) = \sum_{k=1}^n k^2 \frac{1}{n} = \frac{n(n+1)(2n+1)}{6n} = \frac{(n+1)(2n+1)}{6} \, .
$$

Thus

$$
\begin{aligned}
Var(X) &= E(X^2) - (EX)^2 \\
&= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\
&= \frac{n+1}{2} \left\{ \frac{(2n+1)}{3} - \frac{n+1}{2} \right\} \\
&= \frac{(n-1)(n+1)}{12} \, .
\end{aligned}
$$

Note that this random variable $X$ has a discrete uniform distribution on $\{1, \ldots, n\}$, and hence the above mean and variance are comparable to those of $Y \equiv nU$ where $U \sim \text{Uniform}(0,1)$, namely $E(Y) = n/2$ and $Var(Y) = n^2/12$. Put another way, $X/n$ has a discrete uniform distribution on $\{1/n, \ldots, n/n = 1\}$, and $E(X/n) = (n+1)/(2n) \to 1/2 = E(U)$, $Var(X/n) = (n-1)(n+1)/(12n^2) \to 1/12 = Var(U)$.

7. K, 4.3, # 3, page 269. Find $EX$, $EX^2$, and $Var(X)$ if $X$ has the density $f(x) = \lambda^2 x e^{-\lambda x} 1_{(0,\infty)}(x)$.
(Note that there are two way to do this: (i) direct calculation; (ii) recognize the given density as that of a sum of two independent random variables with a distribution we have already studied.)

**Solution:** Solution 1 (direct calculation). For any positive integer $k$

$$\begin{aligned} E(X^k) &= \int_0^\infty x^k f(x)dx = \int_0^\infty x^k \lambda^2 x e^{-\lambda x} dx \\ &= \int_0^\infty x^{k+1} \lambda^2 e^{-\lambda x} dx = \frac{1}{\lambda^k} \int_0^\infty y^{k+1} e^{-y} dy \\ &= \frac{(k+1)!}{\lambda^k} \int_0^\infty \frac{1}{(k+1)!} y^{k+1} e^{-y} dy \\ &= \frac{(k+1)!}{\lambda^k} \cdot 1 = \frac{(k+1)!}{\lambda^k} \end{aligned}$$

since $y^{k+1}e^{-y}/(k+1)!$ is the Gamma$(k+2, 1)$ density which integrates to 1. In particular, $E(X) = 2/\lambda$ and $E(X^2) = 6/\lambda^2$, so that $Var(X) = E(X^2) - (E(X))^2 = (6-4)/\lambda^2 = 2/\lambda^2$.
Solution 2 (recognition of $X$ as a sum of independent rv's). Note that the density $f$ is a Gamma$(2, \lambda)$ density, i.e. the density of the waiting time until the second event in a Poisson process with intensity parameter $\lambda$. Thus $X = Y_1 + Y_2$ where $Y_i \sim$ Exponential$(\lambda)$. Now $E(Y_i) = 1/\lambda$ and $Var(Y_i) = 1/\lambda^2$ by our earlier calculations. Hence by Fact 2

$$E(X) = E(Y_1 + Y_2) = E(Y_1) + E(Y_2) = 2/\lambda, ,$$

and by Fact 4 and using independence of $Y_1$ and $Y_2$,

$$Var(X) = Var(Y_1) + Var(Y_2) = \lambda^{-2} + \lambda^{-2} = 2/\lambda^2 .$$

8. Bonus Problem 1: K, 4.2, # 9, page 257. Let $X$ be the number of failures before the first success in Bernoulli trials with success probability $p$. That is, $X$ has the geometric distribution with parameter $p$; the mass function is $p(k) = pq^k$ for $k = 0, 1, \ldots$. In example (4.1.3) we found that $E(X) = q/p$.
(a) Show that $E[X(X-1)] = 2q^2/p^2$.
(b) From (a) and $EX$ find $E(X^2)$.
(c) Use the result of Exercise 4 above with $EX$ and $E(X^2)$ to find $Var(X) = E(X - EX)^2$.

**Solution:** First note that $Y \equiv X + 1$ is the number of *trials* until the first success, and hence $Y$ is the random variable which has a Geometric$(p)$ distribution according to our Bernoulli process terminology: the mass function of $Y$ is $p_Y(y) = q^{y-1}p$ for $y = 1, 2, \ldots$. In our discussion of the Geometric and Negative Binomial distributions in Handout 3 we showed that $E(Y) = 1/p$ and $Var(Y) = q/p^2$. This implies that $E(X) = E(Y) - 1 = 1/p - 1 = q/p$ and $Var(X) = q/p^2$. Thus

$$E(X^2) = Var(X) + (EX)^2 = q/p^2 + (q/p)^2 = (q/p^2)(1+q) = q(2-p)/p .$$

Here is the computation by direct calculation: (a)

$$\begin{aligned} E[X(X-1)] &= \sum_{k=0}^{\infty} k(k-1)q^k p = pq^2 \sum_{k=2}^{\infty} k(k-1)q^{k-2} \\ &= pq^2 \frac{d^2}{dq^2} \sum_{k=0}^{\infty} q^k = pq^2 \frac{d^2}{dq^2} \frac{1}{1-q} \\ &= pq^2 \frac{d}{dq} \frac{1}{(1-q)^2} = pq^2 \frac{2}{(1-q)^3} = 2q^2/p^2 \,. \end{aligned}$$

Thus we have

$$(q/p)^2 = E(X^2) - E(X) = E(X^2) - q/p$$

and hence

$$E(X^2) = 2(q/p)^2 + q/p = (q/p)[2(q/p)+1] = q(2-p)/p^2 \,.$$

This yields

$$Var(X) = E(X^2) - (EX)^2 = q(2-p)/p^2 - (q/p)^2 = q/p^2$$

which agrees with our previous calculation.

9. Bonus Problem 2: Look at the Ball and Urn Experiment at the Virtual Laboratories website:

   `http://www.math.uah.edu/stat/urn/index.html`

   (a) Find a formula for the variance of the total number of red balls in a sample of size $n$ drawn without replacement from an urn containing $R$ red balls and a total of $N$ balls.
   (b) Find the corresponding variance when the sampling is carried out *with* replacement.
   (c) For the sampling without replacement part of this experiment, with $N = 60$ and $R = 20$, verify the claim that the standard deviation of $Y$, the number of red balls in $n = 10$ draws from the urn, is 1.37.

   **Solution:** (a). As we have discussed in class and in Handout # 9, if $T_n \sim$ Hypergeometric$(R, N, n)$, then

   $$Var(T_n) = n\frac{R}{N}(1-\frac{R}{N})\left(1-\frac{n-1}{N-1}\right) \,.$$

   (b) The corresponding variance for sampling *with replacement* is

   $$Var(T_n) = n\frac{R}{N}(1-\frac{R}{N}) \,.$$

   (c) When $N = 60$, $R = 20$, and $n = 10$ we compute

   $$\begin{aligned} Var(T_10) &= 10\frac{20}{60}(1-\frac{20}{60})\left(1-\frac{9}{59}\right) \\ &= 1.883239 \,. \end{aligned}$$

   Hence $\sigma_{T_{10}} = 1.372312$, in agreement with the claimed standard deviation.

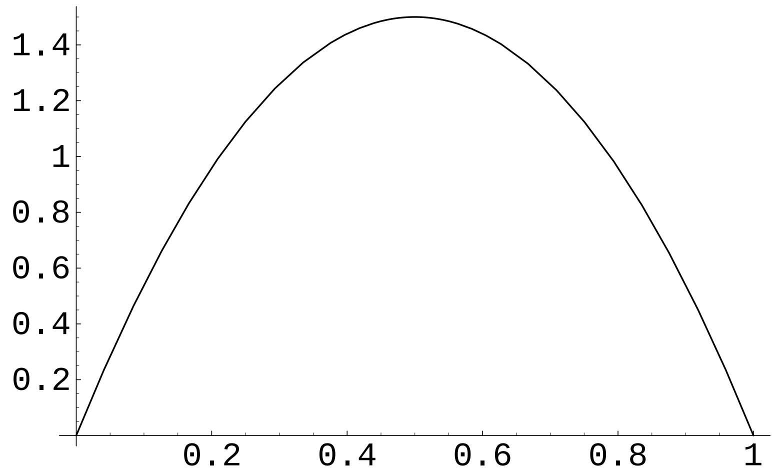


Figure 1: The density $f(x) = 6x(1-x)1_{[0,1]}(x)$.

10. Bonus Problem 3: For each of the following problems on Problem Set 7, compute the variance of the random variables with the given distributions and sketch the mass or density function: problem # 2 (K 4.1, #3); problem # 3 (K 4.1, #5); problem # 4 (K 4.1, #7).

**Solution:** Problem # 2 (K 4.1, #3).
(a) If $f(x) = 6x(1-x)1_{(0,1)}(x)$, then we calculated $E(X) = 1/2$, and

$$E(X^2) = \int_0^1 x^2 6x(1-x)dx = 6(\int_0^1 x^3 dx - \int_0^1 x^4 dx) = 6(1/4-1/5) = 6/20 = 3/10\,,$$

and hence $Var(X) = E(X^2) - (EX)^2 = 3/10 - 1/4 = (12-10)/40 = 1/20$.
See Figure 1 for a plot of the density function $f$.

(b) If $f(x) = (3/x^4)1_{(1,\infty)}(x)$, then

$$E(X^2) = \int_1^\infty x^2(3/x^4) = \int_1^\infty (3/x^2)dx = 3\,,$$

so

$$Var(X) = E(X^2) - (EX)^2 = 3 - (3/2)^2 = 12/4 - 9/4 = 3/4\,.$$

See Figure 2 for a plot of the density $f$.

Problem # 3 (K 4.1, #5):
This is just sampling without replacement from an urn containing the numbers

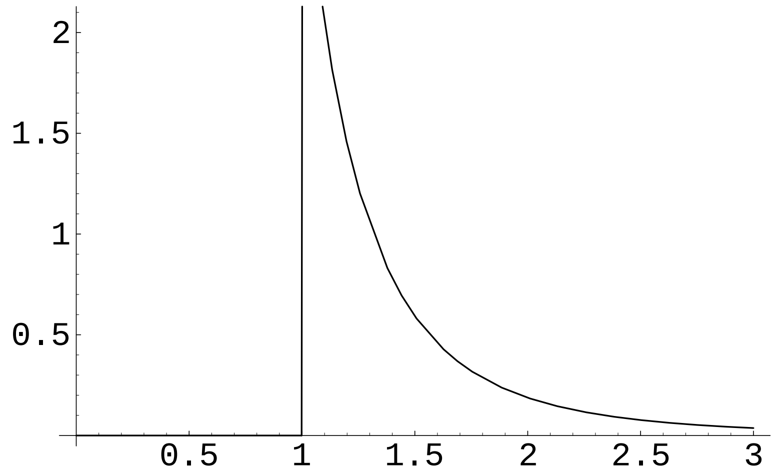


Figure 2: The density $f(x) = (3/x^4)1_{[1,\infty)}(x)$.

$\{a_1, \ldots, a_5\} = \{1, \ldots, 5\}$. Thus $N = 5$, $n = 2$, and Thus $X = T_n = T_2$. From the development in Handout #9,

$$Var(T_2) = \left(1 - \frac{2-1}{5-1}\right) 2\sigma_a^2 = \left(1 - \frac{1}{4}\right) \cdot 2 \cdot 2 = 3\,.$$

since

$$\begin{aligned} \sigma_a^2 &\equiv \frac{1}{N}\sum_{i=1}^{N}(a_i - \overline{a}_N)^2 = \frac{1}{5}\sum_{i=1}^{5} i^2 - 3^2 \\ &= \frac{1}{5}\frac{5(5+1)(2\cdot 5+1)}{6} - 3^2 = \frac{6 \cdot 11}{6} - 9 = 2\,. \end{aligned}$$

Note that this agrees with the result in problem 6 above: for $n = 5$, $(n-1)(n+1)/12 = 4 \cdot 6/12 = 2$.
Problem # 4 (K 4.1, #7):
Since $p(k) = (k-1)(1/2)^k$ for $k = 2, 3, \ldots$, we know (as mentioned in the solution for problem set 7) that $X$ has the same distribution as $W_2 \sim$ NegativeBinomial$(2, 1/2)$. Thus $X = W_2 = Y_1 + Y_2$ where $Y_i \sim$ Geometric$(p)$. Hence

$$Var(X) = Var(Y_1) + Var(Y_2) = 2 \cdot (1/2)/(1/2)^2 = 4\,.$$

See Figure 3 for a plot of the probability mass function $p$.

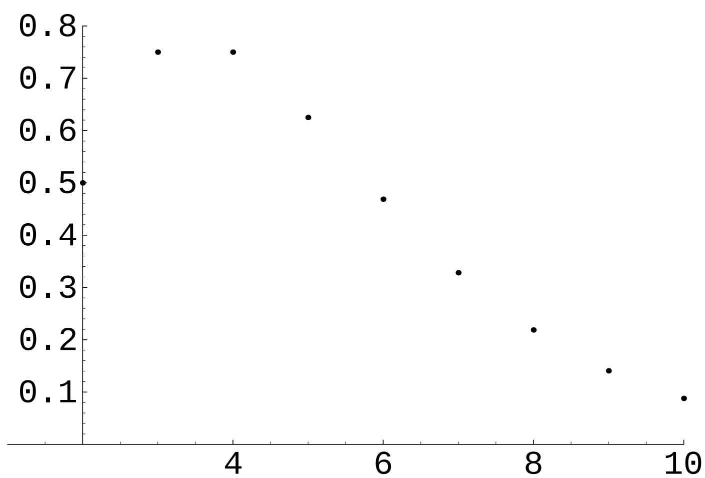


Figure 3: The mass function $p(k) = k(k-1)(1/2)^k 1_{\{2,3,\ldots\}}(k)$.