STA2001 Assignment 3

Due Date: June 27, 2023

(3.1-3). Customers arrive randomly at a bank teller's window. Given that one customer arrived during a particular 10-minute period, let X equal the time within the 10 minutes that the customer arrived. If X is $U(0,10)$, find
(a) The pdf of X.
(b) $P(X \geq 8)$.
(c) $P(2 \leq X < 8)$.
(d) $E(X)$.
(e) $\text{Var}(X)$.

Solution:

(a) The pdf is $f(x) = \frac{1}{10}, \quad 0 < x < 10$

(b) $P(X \geq 8) = \int_8^{10} \frac{1}{10} dx = 0.2$

(c) $P(2 \leq X < 8) = \int_2^8 \frac{1}{10} dx = 0.6$

(d) $E(X) = \frac{0+10}{2} = 5$

(e) $\text{Var}(X) = \frac{(10-0)^2}{12} = \frac{25}{3}$

(3.1-5). Let Y have a uniform distribution $U(0,1)$, and let

$$W = a + (b-a)Y, \quad a < b.$$

(a) Find the cdf of W.
Hint: Find $P[a + (b-a)Y \leq w]$.
(b) How is W distributed?

Solution:

(a) Note that pdf of Y is just 1. Let G be the cdf of W, then for $a < w < b$

$$G(w) = P(W \leq w) = P(a + (b-a)Y \leq w) = P\left(Y \leq \frac{w-a}{b-a}\right) = \int_0^{\frac{w-a}{b-a}} 1\, dy = \frac{w-a}{b-a}$$

and $G(w) = 0$ for $w \leq a$ and $G(w) = 1$ for $w \geq b$.

(b) From the cdf of W in (a), we know that $W \sim U(a,b)$.

(3.1-6). A grocery store has n watermelons to sell and makes \$1.00 on each sale. Say the number of consumers of these watermelons is a random variable with a distribution that can be approximated by

$$f(x) = \frac{1}{200}, \quad 0 < x < 200,$$

a pdf of the continuous type. If the grocer does not have enough watermelons to sell to all consumers, she figures that she loses \$5.00 in goodwill from each unhappy customer. But if she has surplus watermelons, she loses 50 cents on each extra watermelon. What should n be to maximize profit?

Hint: If $X \leq n$, then her profit is $(1.00)X + (-0.50)(n - X)$; but if $X > n$, her profit is $(1.00)n + (-5.00)(X - n)$. Find the expected value of profit as a function of n, and then select n to maximize that function.

Solution:

From the hint, we let $g(X)$ be the profit function and write

$$g(X) = \begin{cases} (1.00)X + (-0.50)(n - X), & 0 \leq X \leq n \\ (1.00)n + (-5.00)(X - n), & 200 \geq X > n \end{cases}$$

where we have used the fact that the continuous random variable X only has positive density on $(0, 200)$.

So the mean of $g(X)$can be further written as

$$\begin{aligned} E(g(X)) &= \int_0^{200} g(x)f(x)dx \\ &= \left(\int_0^n (1.5x - 0.5n)\, dx + \int_n^{200} (-5x + 6n)\, dx\right) \cdot \frac{1}{200} \\ &= \frac{1}{200}\left(-\frac{13}{4}n^2 + 1200n - 100000\right) \end{aligned}$$

Using the first order condition (i.e. set the first derivative as zero),

$$\frac{d}{dn}E(g(X)) = \frac{d}{dn}\frac{1}{200}\left(-\frac{13}{4}n^2 + 1200n - 100000\right) = \frac{1}{200}\left(-\frac{13}{2}n + 1200\right) = 0$$

and solve the equation we yields

$$n = \frac{2400}{13} = 184.6153 \approx 185$$

Note that we could check that it's indeed a maximum point as the function $E(g(X))$ is concave in n, i.e. its second derivative is negative.

(3.2-1). What are the pdf, the mean, and the variance of X if the moment-generating function of X is given by the following?
(a) $M(t) = \frac{1}{1-3t}, t < 1/3$.
(b) $M(t) = \frac{3}{3-t}, t < 3$.

Solution:

(a) By matching the moment generating function, we know that X has an exponential distribution with parameter $\theta = 3$. The corresponding pdf is given by

$$f(x) = \frac{1}{3}e^{-x/3}, \quad 0 \leq x < \infty$$

and zero elsewhere.

Since it's an exponential distribution, its mean and variance are

$$E(\mathrm{X}) = \theta = 3, \qquad \mathrm{Var}(X) = \theta^2 = 9$$

(b) By matching the moment generating function, we know that X has an exponential distribution with parameter $\theta = 1/3$. The corresponding pdf is given by

$$f(x) = 3e^{-3x}, \quad 0 \leq x < \infty$$

and zero elsewhere.

Since it's an exponential distribution, its mean and variance are

$$E(X) = \theta = \frac{1}{3}, \qquad \mathrm{Var}(X) = \theta^2 = \frac{1}{9}$$

Note: we could also compute the moments (and thus the variance) by using the property $E(X^k) = M^{(k)}(t)|_{t=0}$, for positive integers k and a given mgf. This is a general method and it's important in our course.

(3.2-3). Let X have an exponential distribution with mean $\theta > 0$. Show that

$$P(X > x + y|X > x) = P(X > y)$$

for any $x > 0$.

Solution:

For $x > 0$ and $y < 0$, it's obvious that $P(X > x + y|X > x) = 1 = P(X > y)$, as the cdf of an exponential distribution says that $P(X > y) = 1$ for any $y < 0$.

Now we consider $x > 0$ and $y > 0$. By Bayes' theorem,

$$P(X > x + y|X > x) = \frac{P(X > x + y)}{P(X > x)} = \frac{e^{-(x+y)/\theta}}{e^{-x\theta}} = e^{-y\theta} = P(X > y)$$

which completes the proof.

(3.2-7). Find the moment-generating function for the gamma distribution with parameters α and θ. Hint: In the integral representing $E\left(e^{tX}\right)$, change variables by letting $y = (1-\theta t)x/\theta$, where $1-\theta t > 0$.

Solution: Let $f(x)$ be the pdf of the given Gamma distribution, then by definition of the MGF,

$$\begin{aligned}
M(t) &= E\left[e^{tX}\right] \\
&= \int_0^{+\infty} e^{tx} f(x)dx \\
&= \int_0^{+\infty} e^{tx} \cdot \frac{x^{\alpha-1}e^{-x/\theta}}{\Gamma(\alpha)\theta^\alpha}dx \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha}\int_0^{+\infty} x^{\alpha-1}e^{-x/\theta}e^{tx}dx \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha}\int_0^{+\infty} x^{\alpha-1}e^{-x(1-\theta t)/\theta}dx \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha}\int_0^{+\infty} \left(\frac{\theta u}{1-\theta t}\right)^{\alpha-1} e^{-u}\frac{\theta}{(1-\theta t)}du \\
&= \frac{\theta^\alpha}{\Gamma(\alpha)\theta^\alpha(1-\theta t)^\alpha}\int_0^{+\infty} u^{\alpha-1}e^{-u}du \\
&= \frac{\theta^\alpha}{\Gamma(\alpha)\theta^\alpha(1-\theta t)^\alpha}\Gamma(\alpha) \\
&= \frac{1}{(1-\theta t)^\alpha}
\end{aligned}$$

where we let $u = (1-\theta t)x/\theta$ in the 6th equality and thus $du = \frac{(1-\theta t)}{\theta}dx$. The final results is obtained by noticing that $\Gamma(\alpha) := \int_0^{+\infty} x^{\alpha-1}e^{-x}dx$.

(3.2-11). If X is $\chi^2(17)$, find

(a) $P(X < 7.564)$
(b) $P(X > 27.59)$
(c) $P(6.408 < X < 27.59)$
(d) $\chi^2_{0.95}(17)$
(e) $\chi^2_{0.025}(17)$

Solution:

Since X is $\chi^2(17)$ and degree of freedom is $r = 17$, we can check the chi-square distribution table and find that

(a) $P(X < 7.564) = 0.025$

(b) $P(X > 27.59) = 1 - P(X \leq 27.59) = 1 - 0.95 = 0.05$

(c) $P(6.408 < X < 27.59) = P(X < 27.59) - P(X < 6.408) = 0.95 - 0.01 = 0.94$

(d) $\chi^2_{0.95}(17) = 8.672$

(e) $\chi^2_{0.025}(17) = 30.19$

(3.2-22). Let X have a logistic distribution with pdf

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \quad -\infty < x < \infty.$$

Show that

$$Y = \frac{1}{1+e^{-X}}$$

has a $U(0,1)$ distribution.

Hint: Find $G(y) = P(Y \le y) = P\left(\frac{1}{1+e^{-X}} \le y\right)$, where $0 < y < 1$.

Solution:

We aim to use the the cdf of X to compute the cdf of Y according to the hint. So we compute the cdf of X first, which is obtained by

$$F_X(x) = \int_{-\infty}^{x} \frac{e^{-w}}{(1+e^{-w})^2}\, dw = \frac{1}{1+e^{-x}}, \quad -\infty < x < \infty$$

Note that by definition of Y and $-\infty < x < \infty$, we must have $0 < Y < 1$.

$$\begin{aligned} G(y) &= P\left(\frac{1}{1+e^{-X}} \le y\right) \\ &= P\left(1+e^{-X} \ge \frac{1}{y}\right) \\ &= P\left(X \le -\ln\left(\frac{1}{y}-1\right)\right) \\ &= F_X\left(-\ln\left(\frac{1}{y}-1\right)\right) \\ &= \frac{1}{1+e^{\ln(1/y-1)}} \\ &= y \end{aligned}$$

where $0 < y < 1$. This shows that $Y \sim U(0,1)$.

(3.3-10). If X is $N\left(\mu, \sigma^2\right)$, show that the distribution of $Y = aX + b$ is $N\left(a\mu + b, a^2\sigma^2\right)$ $a \neq 0$. Hint: Find the cdf $P(Y \leq y)$ of Y, and in the resulting integral, let $w = ax + b$ or, equivalently, $x = (w - b)/a$.

Solution:

Let's first consider the case $a > 0$.

$$
\begin{aligned}
G(y) &= P(Y \leq y) \\
&= P(aX + b \leq y) \\
&= P\left(X \leq \frac{y-b}{a}\right) \\
&= \int_{-\infty}^{(y-b)/a} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\, dx \\
&= \int_{-\infty}^{y} \frac{1}{a\sigma\sqrt{2\pi}} e^{-(w-b-a\mu)^2/2\sigma^2 a^2}\, dw
\end{aligned}
$$

where we let $w = ax + b$ in the last equality and thus $dw = adx$.

For $a < 0$, the derivation is quite similar, except that the standard deviation should be $-a\sigma$.

The final result shows that it is indeed a normal cumulative distribution function of $N\left(b + a\mu, \sigma^2 a^2\right)$.

(3.3-14). The strength X of a certain material is such that its distribution is found by $X = e^Y$, where Y is $N(10, 1)$. Find the cdf and pdf of X, and compute $P(10,000 < X < 20,000)$.

Note: $F(x) = P(X \leq x) = P\left(e^Y \leq x\right) = P(Y \leq \ln x)$ so that the random variable X is said to have a lognormal distribution.

Solution:

WLOG, let $Y \sim N(\mu, \sigma^2)$, then the pdf and cdf of Y are given by

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2} \\ F_Y(y) &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-\mu)^2/2\sigma^2} dz \end{aligned}$$

Then, the cdf of X can be computed by

$$F_X(x) = P(X \leq x) = P\left(e^Y \leq x\right) = P(Y \leq \ln x) = F_Y(\ln x) = \int_{-\infty}^{\ln x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-\mu)^2/2\sigma^2} dz$$

The pdf of X is the derivative of the cdf of X,

To compute $P(10000 < X < 20000)$, we write

$$\begin{aligned} P(10000 < X < 20000) &= P(X \leq 20000) - P(X \leq 10000) \\ &= P\left(e^Y \leq 20000\right) - P\left(e^Y \leq 10000\right) \\ &= P(Y \leq \ln 20000) - P(Y \leq \ln 10000) \\ &= P(Z < -0.10) - P(Z < -0.79) \\ &= 1 - P(Z < 0.10) - (1 - P(Z < 0.79)) \\ &= 1 - 0.5398 - (1 - 0.7852) \\ &= 0.2454 \end{aligned}$$

where we have used the fact $Z = \dfrac{Y-\mu}{\sigma} \sim N(0, 1)$.

$$\begin{aligned} f_X(x) &= F_X'(x) \\ &= \frac{d}{dx}\left(\int_{-\infty}^{\ln x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-\mu)^2/2\sigma^2} dz\right) \\ &= \frac{d}{dx}\left(\int_{0}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(\ln t-\mu)^2/2\sigma^2} \frac{1}{t}\, dt\right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(\ln x-\mu)^2/2\sigma^2} \left(\frac{1}{x}\right) \\ &= \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-(\ln x-\mu)^2/2\sigma^2} \end{aligned}$$

where we have used $t = e^z$ in the 3rd equality and thus $dz = 1/u \cdot du$, and the derivative is obtained by the Fundamental theorem of calculus (or you can finish these two steps in one trial by using the Leibniz integral rule).

Finally we set $\mu = 10$ and $\sigma = 1$,

$$\begin{aligned} F_X(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\ln x} e^{-(z-10)^2/2} dz \\ f_X(x) &= \frac{1}{x\sqrt{2\pi}} e^{-(\ln x-10)^2/2} \end{aligned}$$