STAT310

Practice Problems

Week 5

March 26, 2012

1 Continuous random variables.

1. $X_1 \sim$ Uniform(0,20). In this case, the expected value of the distance the pebble lands away from me on the first skip is

$$\begin{aligned} E[X_1] &= \frac{(20+0)}{2} \\ &= 10 \end{aligned}$$

Also, we can see that the probability that the pebble will land between 10-20 feet away from me is equal to

$$\begin{aligned} P(10 < X_1 \leq 20) &= \int_{10}^{20} \frac{1}{20} dx \\ &= \frac{1}{20} [x]_{10}^{20} \\ &= 0.5 \end{aligned}$$

2. $X_2 \sim \mathcal{N}(\mu, \frac{\sigma^2}{n}) = \mathcal{N}(10, \frac{100/3}{100}) = \mathcal{N}(10, \frac{1}{3})$ (since $E[X_1] = 10$ and $Var[X_1] = \frac{(20-10)^2}{12} = \frac{100}{3}$). Thus the probability that the mean distance the pebble lands from me will be between 10-20 feet is equal to

$$\begin{aligned} P(10 < X_2 \leq 20) &= \int_{10}^{20} \frac{1}{\sqrt{2 \cdot \frac{1}{3} \cdot \pi}} e^{-(x-10)^2/(2 \cdot \frac{1}{3})} \\ &\approx 0.5 \end{aligned}$$

3. $X_3 \sim gamma(\alpha = 2.5, \beta = 2)$. The gamma distribution is a natural distribution for describing the precision, because $\sigma^2 \in [0, \infty)$, and the support of a gamma random variable is also over $[0, \infty)$. Note that we can calculate the parameters α, β for this family since

$$\left\{ \begin{aligned} E[X_3] = \alpha\beta = 5 \\ Var[X_3] = \alpha\beta^2 = 10 \end{aligned} \right\}$$
$$\implies \beta = 2, \alpha = 2.5.$$

Thus the probability that the precision is between 2 and 3 is equal to

$$\begin{aligned} P(2 < X_3 \leq 3) &= \int_{2}^{3} \frac{1}{\Gamma(2.5) 2^{2.5}} x^{2.5-1} e^{-x/2} \\ &\approx 0.0597. \end{aligned}$$

4. $X_4 \sim$ exponential(300). The exponential distribution is commonly used to model lifetimes, analogous to the geometric distribution in the discrete case. In fact, both the exponential and geometric distributions share the "memoryless" property, i.e $\forall s > t \geq 0, P(X > s|X > t) = P(X > s - t)$. The variance of X_4 can be thus calculated if one recalls that, for $Y \sim$ exponential(β), then $E[X] = \beta$ and $Var[X] = \beta^2$. Thus
$$Var[X_4] = 300^2 = 90000$$
The probability that the next lightbulb is a "dud" is equal to the probability that it expires 1 hour of use, i.e.
$$\begin{aligned} P(X_4 \leq 1) &= \int_0^1 \frac{1}{300} e^{-x/300} \\ &\approx 0.00332. \end{aligned}$$

2 Normal distributions.

Let the random variable X denote the grades in this class, so that $X \sim \mathcal{N}(72, 36)$.

1. We want to calculate x such that $P(X > x) = 0.15$. We have that
$$\begin{aligned} & P(X > x) = 0.15 \\ \implies \quad & P(Z > \frac{x - 72}{6}) = 0.15 \quad (using\ z - transform) \end{aligned}$$
Using a z-table, this gives
$$\begin{aligned} & \frac{x - 72}{6} = 1.04 \\ \implies \quad & x = 78.4. \end{aligned}$$
Hence the minimum score to get an A is 78.4%.
2. 15%
3. 50%
4. 25%

3 Transformations of PDFs.

1. We are given that $X \sim \mathcal{N}(0, 1)$, i.e.
$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in (-\infty, \infty).$$
Consider $Y = g(X) = X^2$. Then the support of Y is $\mathcal{Y} = (0, \infty)$. Because this function $g(x)$ is monotone on $(-\infty, 0)$ and $(0, \infty)$, then take
$$\begin{aligned} & A_0 = \{0\} \\ & A_1 = (-\infty, 0), \quad g_1(x) = x^2, \quad g_1^{-1}(y) = -\sqrt{y} \\ & A_2 = (0, \infty), \quad g_2(x) = x^2, \quad g_2^{-1}(y) = \sqrt{y}. \end{aligned}$$

Then the PDF of Y is given by

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{y})^2/2} \left| -\frac{1}{2\sqrt{y}} \right. + \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} \left| \frac{1}{2\sqrt{y}} \right. \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} e^{-y/2}, \end{aligned}$$

$y \in (0, \infty)$. In other words, Y is a **chi-squared random variable** with 1 df.

2. We have given that $X \sim gamma(n, \beta)$, i.e.

$$f_X(x) = \frac{1}{\Gamma(n)\beta^n} x^{n-1} e^{-x/\beta}, \quad 0 < x < \infty.$$

Consider $Y = g(X) = \frac{1}{X}$. Then the support of Y is $\mathcal{Y} = (0, \infty)$. Let $y = g(x)$. Then $g^{-1}(y) = \frac{1}{y}$ and $\frac{d}{dy} g^{-1}(y) = -\frac{1}{y^2}$. Thus the PDF of Y is given by

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| \\ &= \frac{1}{\Gamma(n)\beta^n} \left(\frac{1}{y} \right)^{n-1} e^{-1/(\beta y)} \frac{1}{y^2} \\ &= \frac{1}{\Gamma(n)\beta^n} \left(\frac{1}{y} \right)^{n+1} e^{-1/(\beta y)}, \end{aligned}$$

$y \in (0, \infty)$. In other words, Y has a **inverse-gamma** PDF, a natural distribution for describing the variance of a normal distribution.

3. For $Y = F_X(X)$ we have, $\forall 0 < y < 1$,

$$\begin{aligned} P(Y \le y) &= P(F_X(X) \le y) \\ &= P(F_X^{-1}[F_X(X)] \le F_X^{-1}(y)) \quad (since\ F_X^{-1}\ is\ increasing) \\ &= P(X \le F_X^{-1}(y)) \\ &= F_X(F_X^{-1}(y)) \quad (definition\ of\ F_X) \\ &= y \quad (continuity\ of\ F_X). \end{aligned}$$

At the endpoints we have $P(Y \le y) = 1$ for $y \ge 1$ and $P(Y \le y) = 0$ for $y \le 0$, showing that $Y \sim uniform[0, 1]$.

Thus we have that, for X with continuous CDF $F_X(x)$, $Y = F_X(X) \sim uniform[0, 1]$. This important result can be used in simulation to generate random samples from a given distribution. In particular, if we need to generate an observations X from a population with CDF $F_X(x)$, then we need only to generate a uniform random number U from a $uniform[0, 1]$ distribution, and then solve for x in the equation $F_X(x) = u$. This allows us to generate numbers from any distribution!