Probability Review — Diagnostic Problem Set

Graduate Random Processes ECE 534

Time: 40 minutes. Justify every step.

These problems review the undergraduate essentials we will lean on all semester: discrete distributions and indicator methods, cumulative distribution functions and densities (including mixed distributions), transformations of a random variable, expectation, and the Gaussian.

Problem 1 (Discrete random variables; indicators and moments). At a coat check, n guests hand in their hats. On the way out the hats are returned in a uniformly random order, so that each of the $n!$ assignments of hats to guests is equally likely. Let X be the number of guests who receive their own hat.

(a) Find $\mathbb{E}[X]$.

(b) Find $\mathrm{Var}(X)$.

Solution. Write $X = \sum_{i=1}^{n} I_i$, where $I_i = \mathbf{1}\{\text{guest } i \text{ gets their own hat}\}$.

(a) By symmetry guest i is equally likely to receive any of the n hats, so $\mathbb{P}(I_i = 1) = 1/n$ and $\mathbb{E}[I_i] = 1/n$. By linearity of expectation (no independence needed),

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[I_i] = n \cdot \tfrac{1}{n} = 1.$$

(b) Since $I_i \in \{0, 1\}$ we have $I_i^2 = I_i$, so $\mathrm{Var}(I_i) = \frac{1}{n} - \frac{1}{n^2}$. For $i \neq j$,

$$\mathbb{E}[I_i I_j] = \mathbb{P}(I_i = 1, I_j = 1) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}, \qquad \mathrm{Cov}(I_i, I_j) = \frac{1}{n(n-1)} - \frac{1}{n^2}.$$

There are n variance terms and $n(n-1)$ ordered off-diagonal pairs, so

$$\mathrm{Var}(X) = \sum_{i} \mathrm{Var}(I_i) + \sum_{i \neq j} \mathrm{Cov}(I_i, I_j) = n\big(\tfrac{1}{n} - \tfrac{1}{n^2}\big) + n(n-1)\Big(\tfrac{1}{n(n-1)} - \tfrac{1}{n^2}\Big).$$

The first bracket is $1 - \frac{1}{n}$ and the second is $1 - \frac{n-1}{n} = \frac{1}{n}$, so

$$\mathrm{Var}(X) = \Big(1 - \tfrac{1}{n}\Big) + \tfrac{1}{n} = 1.$$

Problem 2 (CDFs and densities; a continuous random variable). A continuous random variable X has density

$$f_X(x) = \begin{cases} c\,x^2, & 0 \le x \le 2, \\ 0, & \text{otherwise}, \end{cases}$$

for some constant c.

(a) Find c and the CDF F_X.

(b) Find the median of X, and compute $\mathbb{P}\big(X > 1 \mid X > \frac{1}{2}\big)$.

(c) Compute $\mathbb{E}[X]$ and $\operatorname{Var}(X)$.

Solution. *(a)* Normalization requires $\int_0^2 c\,x^2\,\mathrm{d}x = c \cdot \frac{8}{3} = 1$, so $c = \frac{3}{8}$. Integrating the density,

$$F_X(x) = \int_0^x \tfrac{3}{8}t^2\,\mathrm{d}t = \frac{x^3}{8}, \qquad 0 \le x \le 2,$$

with $F_X(x) = 0$ for $x < 0$ and $F_X(x) = 1$ for $x > 2$.

(b) The median m solves $F_X(m) = \frac{m^3}{8} = \frac{1}{2}$, i.e. $m^3 = 4$, so $m = 4^{1/3} = 2^{2/3} \approx 1.587$. For the conditional probability, note $\{X > 1\} \subseteq \{X > \frac{1}{2}\}$, so

$$\mathbb{P}\big(X > 1 \mid X > \tfrac{1}{2}\big) = \frac{\mathbb{P}(X > 1)}{\mathbb{P}(X > \frac{1}{2})} = \frac{1 - \frac{1}{8}}{1 - \frac{(1/2)^3}{8}} = \frac{7/8}{63/64} = \frac{8}{9}.$$

(c)

$$\mathbb{E}[X] = \int_0^2 x \cdot \tfrac{3}{8}x^2\,\mathrm{d}x = \tfrac{3}{8} \cdot \frac{2^4}{4} = \frac{3}{2}, \qquad \mathbb{E}[X^2] = \int_0^2 x^2 \cdot \tfrac{3}{8}x^2\,\mathrm{d}x = \tfrac{3}{8} \cdot \frac{2^5}{5} = \frac{12}{5}.$$

Hence $\operatorname{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \dfrac{12}{5} - \dfrac{9}{4} = \dfrac{48 - 45}{20} = \dfrac{3}{20}$.

Problem 3 (Function of a random variable; non-monotone map). Let $X \sim \text{Uniform}(-1, 2)$ and define $Y = X^2$.

(a) Find the density f_Y on its full support. Be careful: $g(x) = x^2$ is not one-to-one on $[-1, 2]$.

(b) Verify that f_Y integrates to 1, and explain the discontinuity in f_Y.

Solution. Here $f_X(x) = \frac{1}{3}$ on $[-1, 2]$, and $Y = X^2$ takes values in $[0, 4]$.

(a) Work through the CDF. For $y \geq 0$, $F_Y(y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y})$.

- If $0 \leq y \leq 1$ then $\sqrt{y} \leq 1$, so $[-\sqrt{y}, \sqrt{y}] \subseteq [-1, 2]$ and $F_Y(y) = \frac{1}{3}\big(\sqrt{y} - (-\sqrt{y})\big) = \frac{2\sqrt{y}}{3}$.
- If $1 < y \leq 4$ then $-\sqrt{y} < -1$, so the interval is clipped on the left to -1 and $F_Y(y) = \frac{1}{3}\big(\sqrt{y} - (-1)\big) = \frac{\sqrt{y}+1}{3}$.

Differentiating,

$$f_Y(y) = \begin{cases} \dfrac{1}{3\sqrt{y}}, & 0 < y < 1, \\ \dfrac{1}{6\sqrt{y}}, & 1 < y < 4, \\ 0, & \text{otherwise.} \end{cases}$$

Equivalently, via $f_Y(y) = \sum_{x:\, x^2 = y} f_X(x)/|g'(x)|$ with $|g'(x)| = 2|x|$: for $0 < y < 1$ both preimages $\pm\sqrt{y}$ lie in $[-1, 2]$, giving $\frac{1/3}{2\sqrt{y}} + \frac{1/3}{2\sqrt{y}} = \frac{1}{3\sqrt{y}}$; for $1 < y < 4$ only $+\sqrt{y}$ lies in the support, giving $\frac{1/3}{2\sqrt{y}} = \frac{1}{6\sqrt{y}}$.

(b) Integrating,

$$\int_0^1 \frac{\mathrm{d}y}{3\sqrt{y}} + \int_1^4 \frac{\mathrm{d}y}{6\sqrt{y}} = \frac{1}{3}\big[2\sqrt{y}\big]_0^1 + \frac{1}{6}\big[2\sqrt{y}\big]_1^4 = \frac{2}{3} + \frac{1}{3} = 1.$$

The density drops by a factor of two at $y = 1$ (from $\frac{1}{3}$ to $\frac{1}{6}$ if we compare $f_Y(1^-)$ and $f_Y(1^+)$): for $y < 1$ two branches of X map into y, while for $y > 1$ only the positive branch remains in the support of X.

Problem 4 (Gaussians and the Q function). For $Z \sim N(0,1)$, the standard Gaussian tail probability is

$$Q(x) = \mathbb{P}(Z > x) = \int_x^\infty \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, \mathrm{d}t.$$

(a) Using the symmetry of the standard normal density, show that $Q(-x) = 1 - Q(x)$ (in particular $Q(0) = \frac{1}{2}$). Then show that for $X \sim N(\mu, \sigma^2)$ and any real a, $\mathbb{P}(X > a) = Q\big(\frac{a-\mu}{\sigma}\big)$.

(b) Let $X \sim N(2,4)$. Express in terms of Q: (i) $\mathbb{P}(X > 5)$; (ii) $\mathbb{P}(0 < X < 3)$; (iii) $\mathbb{P}(|X-2| > 3)$.

(c) *(Antipodal signaling.)* A transmitter sends $S = +A$ or $S = -A$, each with probability $\frac{1}{2}$ ($A > 0$). The receiver observes $Y = S + N$, where $N \sim N(0, \sigma^2)$ is independent of S, and decides $\hat{S} = +A$ if $Y > 0$ and $\hat{S} = -A$ otherwise. Find the probability of error $\mathbb{P}(\hat{S} \neq S)$ in terms of Q.

Solution. *(a)* Because Z and $-Z$ have the same distribution,

$$Q(-x) = \mathbb{P}(Z > -x) = \mathbb{P}(-Z > -x) = \mathbb{P}(Z < x) = 1 - \mathbb{P}(Z > x) = 1 - Q(x),$$

where $\mathbb{P}(Z < x) = 1 - \mathbb{P}(Z > x)$ uses continuity ($\mathbb{P}(Z = x) = 0$). Setting $x = 0$ gives $Q(0) = 1 - Q(0)$, so $Q(0) = \frac{1}{2}$. If $X \sim N(\mu, \sigma^2)$ then $Z := (X - \mu)/\sigma \sim N(0,1)$, hence

$$\mathbb{P}(X > a) = \mathbb{P}\Big(Z > \tfrac{a-\mu}{\sigma}\Big) = Q\Big(\tfrac{a-\mu}{\sigma}\Big).$$

(b) Here $\mu = 2$, $\sigma = 2$.

(i) $\mathbb{P}(X > 5) = Q\big(\frac{5-2}{2}\big) = Q(1.5) \approx 0.067$.

(ii) $\mathbb{P}(0 < X < 3) = \mathbb{P}(X > 0) - \mathbb{P}(X > 3) = Q\big(\frac{0-2}{2}\big) - Q\big(\frac{3-2}{2}\big) = Q(-1) - Q(0.5) = \big(1 - Q(1)\big) - Q(0.5) \approx 0.841 - 0.309 = 0.533$.

(iii) Since $(X-2)/2 \sim N(0,1)$, $\mathbb{P}(|X-2| > 3) = \mathbb{P}\big(|Z| > \frac{3}{2}\big) = 2Q(1.5) \approx 0.134$.

(c) Condition on the transmitted symbol. Given $S = +A$, $Y = A + N$ and an error occurs iff $Y \leq 0$, i.e. $N \leq -A$; by symmetry of N, $\mathbb{P}(N < -A) = \mathbb{P}(N > A) = Q(A/\sigma)$. Given $S = -A$, $Y = -A + N$ and an error occurs iff $Y > 0$, i.e. $N > A$, again with probability $Q(A/\sigma)$. By total probability,

$$\mathbb{P}(\hat{S} \neq S) = \tfrac{1}{2}\, Q\Big(\tfrac{A}{\sigma}\Big) + \tfrac{1}{2}\, Q\Big(\tfrac{A}{\sigma}\Big) = Q\Big(\tfrac{A}{\sigma}\Big).$$

The error probability depends only on the ratio A/σ (a signal-to-noise measure): larger separation or smaller noise drives it down.