Math 425, Section 7, Winter 2003: Introduction to Probability

Second Mid-Term: Solutions

1. a.) Let $\mathbf{X}$ be a random variable uniformly distributed over the interval $[\alpha, \beta]$, where $-\infty < \alpha < \beta < +\infty$. Calculate the cumulative distribution function, the expectation and the variance of $\mathbf{X}$.

Solution: The *cumulative distribution function* is given by

$$F_{\mathbf{X}}(x) = \int_{-\infty}^{x} f_{\mathbf{X}}(x)\, dx = P(\mathbf{X} \leq x).$$

In our case this becomes

$$F_{\mathbf{X}}(x) = \begin{cases} 0, & \text{for } -\infty x \leq \alpha, \\ \frac{x-\alpha}{\beta-\alpha}, & \alpha \leq x \leq \beta, \\ 1, & \beta \leq x < +\infty \end{cases}$$

The *expectation* is $E(\mathbf{X}) = \int_{-\infty}^{\infty} x f_{\mathbf{X}}(x)\, dx = \frac{1}{2}(\beta + \alpha)$.

The *variance* is given by

$$\begin{aligned} Var(\mathbf{X}) &= E((\mathbf{X} - \tfrac{\beta-\alpha}{2})^2) \\ &= \tfrac{1}{\beta-\alpha} \int_{\alpha}^{\beta} (x - \tfrac{\beta-\alpha}{2})^2 dx \\ &= \tfrac{(\beta-\alpha)^2}{12}. \end{aligned}$$

b.) Let $\mathbf{Y}$ be a random variable whose probability density function is identically zero except between $x = \alpha$ and $x = \beta$, and on this interval, its graph describes an isosceles triangle of height $\frac{2}{\beta-\alpha}$. Is the variance of $\mathbf{Y}$ greater than, equal to or less than the variance of $\mathbf{X}$ above? Give a reason, but it is *not* required that you calculate the variance of $\mathbf{Y}$ exactly.

Solution: The variance of $\mathbf{Y}$ should be less than the variance of $\mathbf{X}$ because the probability is distributed over the same interval $[\alpha, \beta]$ for both RV's, but the probability density for $\mathbf{Y}$ is more concentrated around the mean than $\mathbf{X}$. The distribution of $\mathbf{X}$ is more diffuse, and hence its variance should be greater.
[In fact, it is not hard to calculate that $Var(\mathbf{Y}) = \frac{(\beta-\alpha)^2}{24}$, half that of $\mathbf{X}$.]

c.) Suppose we had a population of 200 elk, of which a certain number n were tagged. We have to estimate the number of tagged elk by sampling 20 elk. Which method would likely be more efficient to determine this, sampling with replacement or sampling without replacement?

Solution: The agument is similar to part (b), only this time you must recognize that sampling out of 200 elk without replacement is represented by the hypergeometric distribution on the exam sheet, where N would be 200, M would be the number of tagged elk, and n would be the sample size. The variance of this distribution is given as

$$Var = \frac{N-n}{N-1} \cdot n(\frac{M}{N})(1 - \frac{M}{N}).$$

Sampling with replacement would be represented with the binomial distribution with parameters n and $p = \frac{M}{N}$. Hence, its variance is given by

$$Var = n \cdot (\frac{M}{N})(1 - \frac{M}{N}).$$

Comparing the two variances, since for $n > 1$ we have

$$\frac{N-n}{N-1} < 1,$$

the hypergeometric distribution has smaller variance for the same values of N, M, n.

2. Green Pines College would like to have 1050 first year students. They cannot accommodate more than 1060. Assume that each applicant accepts with probabilty 0.6. If the college accepts 1700 students, what is the probability that it will have *too many* acceptances? Be sure to state clearly how you are modeling, what your assumptions are, and what you are using to evaluate this.

Solution: This can be treated with DeMoivre's theorem. We model the situation with a Bernoulli random variable N = "the number of students who accept the offer of admission". It is Bernoulli because we are told that each admitted student accepts *independently* with probability 0.6. The Bernoulli parameters in the problem would be $n = 1700$ trials, with bias $p = 0.6$. We want to estimate the probability $P(\mathbf{N} > 1060)$. To improve accuracy at the edge, we will replace $\mathbf{N} > 1060$ by $\mathbf{N} > 1059.5$, and normalize $\mathbf{N}$ by its mean and variance to compare with the unit normal RV $\mathbf{Z}$. So, we want to estimate

$$P(\frac{\mathbf{N} - 1700 \cdot 0.6}{\sqrt{1700 \cdot 0.6 \cdot (1-0.6)}} > \frac{1059.5 - 1700 \cdot 0.6}{\sqrt{1700 \cdot 0.6 \cdot (1-0.6)}})$$

by $P(\mathbf{Z} > \frac{1059.5-1700 \cdot 0.6}{\sqrt{1700 \cdot 0.6 \cdot (1-0.6)}}) = 1 - \Phi(\frac{1059.5-1700 \cdot 0.6}{\sqrt{1700 \cdot 0.6 \cdot (1-0.6)}}) = 1 - \Phi(1.956)$, which is approximately $1 - 0.9748 = 0.0252$.

3. Let $\mathbf{X}$ be a random variable with probability density function $f_{\mathbf{X}}(x)$. What is the probability density function $f_{\mathbf{Y}}(y)$ of $\mathbf{Y} = a\mathbf{X} + b$?

Solution: This is a direct application of the change of random variable formula:

$$f_{g(\mathbf{X})}(y) = f_{\mathbf{X}}(g^{-1}(y)) \cdot \frac{1}{|g'(g^{-1}(y))|},$$

where $g(x) = ax+b$. If $a = 0$, then $g(\mathbf{X})$ is a constant, and the probability density function is trivial (probabilty 1 at $y = b$). Otherwise, $g^{-1}(y) = \frac{1}{a}(y-b)$, and $g'(g^{-1}(y)) \equiv a$. Putting this together gives

$$f_{g(\mathbf{X})}(y) = \frac{1}{|a|} \cdot f_{\mathbf{X}}(\frac{1}{a}(y-b)).$$

4. A certain typing agency employs two typists. The average number of typing errors per article is different for the two typists: 3 per article for the first and 4.2 for the second. If each of the two typists is equally likely to type your article, what is the approximate probabillity that you will have no mistakes in your typed article. Be sure to be explicit and clear about your modeling.

This uses a Poisson distribution model (for the mistakes, as rare events) for each of the two typists. The Poisson parameters are $\lambda_1 = 3$, and $\lambda_2 = 4.2$. Let A be the event that get the first typist to do the article, and B the event you get the second. Let $\mathbf{N}$ be the RV "the number of typographical errors". Then we can calculate by conditioning:

$$P(\mathbf{N} = 0) = P(\mathbf{N} = 0|A) \cdot P(A) + P(\mathbf{N} = 0|B) \cdot P(B) = \frac{1}{2}e^{-3} + \frac{1}{2}e^{-4.2}.$$

5. Suppose we have a coin whose bias p we do not know, and we flip it 100 times. We observe 55 heads out of these one hundred. We estimate the bias $p = 0.55$. The likelihood of seeing 55 heads is the function $P(\text{observe 55 heads} \mid \text{bias} = p)$. Show that the estimate $p = 0.55$ maximizes the likelihood.

Solution: Since we are talking about Bernoulli trials (I am assuming the flips are *independent*), we know that

$$P(\text{observe 55 heads}|\text{bias} = p) = \binom{100}{55} p^{55}(1-p)^{45}.$$

This function of p is always non-negative, for $p \in [0, 1]$, and is zero for $p = 0$ or 1. So it has a maximum in the interval, where its derivative with respect to p is zero, i.e., where

$$\binom{100}{55} \cdot (55 \cdot p^{54}(1-p)^{45} - 45 \cdot p^{55}(1-p)^{44}) = 0.$$

One solves this equation easily to get $p = 55$.