1. (6p) Two coins are flipped. One of the coins is such that it shows heads with probability $1/3$ whereas the second one shows heads with probability $2/3$. Given that you get exactly one heads, what is the conditional probability that it was the second coin?

Solution: Let $A_i$ be the event that coin number $i$ shows heads. Then

$$\Pr(A_2|A_1 \Delta A_2) = \frac{\Pr(A_2 \cap A_1^c)}{\Pr(A_2 \cap A_1^c) + \Pr(A_2^c \cap A_1)} = \frac{2/3 \cdot 2/3}{2/3 + 2/3} = \frac{4/9}{4/3} = \frac{1}{3}.$$ 

2. (6p) Among twelve light bulbs, there are five defective ones. If one picks uniformly at random four of these light bulbs, what is the probability of getting $x$ defected ones, $x = 0, 1, 2, 3, 4$?

Solution: Let $X$ be the number of defective light bulbs. Then $\Pr(X = k)$ is the ration of the number of ways of picking exactly $k$ defective and $4 - k$ good ones and the total number of ways of picking four bulbs. This gives

$$\Pr(X = k) = \binom{5}{k} \binom{7}{4-k} \binom{12}{4}.$$ 

Computing this for $k = 1, 2, 3, 4$ gives the respective probabilities $7/99, 35/99, 42/99, 14/99, 1/99$.

3. (6p) Suppose that a random number generator uniform numbers among 1, 2, \ldots, 9. What is the probability that the product of $n$ such independent random numbers, is divisible by 10?

Solution: Let $X = X_1X_2 \ldots X_n$. The event that $X$ is not divisible by 10 is the event that either none of the $X_i$’s is divisible by 5 or none of the $X_i$’s is divisible by 2. Let $A$ be the former event and $B$ the latter event. Since the $X_i$’s are independent,

$$\Pr(A) = \Pr(X_1 \text{ not divisible by 5})^n = \left(\frac{8}{9}\right)^n.$$ 

In the same way

$$\Pr(B) = \left(\frac{5}{9}\right)^n.$$ 

Also, the event that $X_1$ is neither divisible 5 nor by 2 is the event that $X_1$ is 1, 3, 7 or 9. Hence $\Pr(A \cap B) = (4/9)^n$. Hence

$$\Pr(A \cup B) = \left(\frac{8}{9}\right)^n + \left(\frac{5}{9}\right)^n - \left(\frac{4}{9}\right)^n.$$ 

Since we are seeking $\Pr((A \cup B)^c)$, the answer is $1 - \left(\frac{8}{9}\right)^n - \left(\frac{5}{9}\right)^n + (4/9)^n$. 

4. (7p) Let $X_1, X_2, \ldots$ be independent and identically distributed random variables with expectation $\mu$. Let $N$ be a positive integer valued random variable such that $\mathbb{E}[N] < \infty$ and such that $I_{\{N \geq n\}}$ is independent of $X_n$ for all $n$. Prove that
\[
\mathbb{E}\left[ \sum_{i=1}^{N} X_i \right] = \mu \mathbb{E}[N].
\]
Solution: We have by the law of total probability that
\[
\mathbb{E}\left[ \sum_{i=1}^{N} X_i \right] = \sum_{n=1}^{\infty} \mathbb{E}\left[ \sum_{i=1}^{n} X_i \mid N = n \right] \mathbb{P}(N = n)
= \sum_{n=1}^{\infty} \mathbb{E}\left[ \sum_{i=1}^{n} X_i \right] \mathbb{P}(N = n)
= \sum_{n=1}^{\infty} \mathbb{E}\left[ \sum_{i=1}^{n} X_i \right] \mathbb{P}(N = n)
= \mu \sum_{n=1}^{\infty} n \mathbb{P}(N = n) = \mu \mathbb{E}[N]
\]
where the third equality follows from that $N$ is independent of the $X_i$'s.

5. (7p) Assume that the traffic at a certain point at a certain road is a Poisson process with intensity $c$. After 30 minutes, you have observed 214 vehicles pass.
(a) Make a maximum likelihood estimate of the intensity $c$,
(b) Use the central limit theorem to give an approximate 95% confidence interval for $c$
Solution: Let $X(t)$ be the number of vehicles that pass up to time $t$. Then $X(t) \sim \text{Poi}(ct)$. We observe $X(30)$ which is Poi(30$c$), so the likelihood is
\[
L(c; x) = f_X(x) = e^{-30c} \frac{(30c)^x}{x!}
\]
so
\[
\ln L(c; x) = -30c + x \ln(30c) - \ln(x!).
\]
Differentiating with respect to $c$ and setting to 0 gives
\[-30 + \frac{x}{c} = 0\]
which gives
\[
\hat{c} = \frac{x}{30}.
\]
Since it was observed that $X(30) = 219$, we get
\[
\hat{c} = \frac{219}{30} = 7.3.
\]
For part (b), the CLT gives that $X(30)$ is approximately $\mathcal{N}(\lambda, \lambda)$ where $\lambda = 30c$. Standardization gives that $(X(30) - \lambda)/\sqrt{\lambda}$ is approximately standard normal. In analogy with what is done for the binomial distribution, we may replace $\lambda$ in the denominator by $\hat{\lambda} = 3\hat{c} = 219$, i.e. $(X(30) - 30c)/\sqrt{219}$ is approximately standard normal. This gives the 95% confidence interval
\[
30c = X(30) \pm 1.96 \sqrt{219}
\]
i.e.
\[
c = \frac{X(30)}{30} \pm 1.96 \frac{\sqrt{219}}{30} = 7.3 \pm 0.97.
\]
6. (6p) In a sample of 15 normally distributed random variables with unknown expectation $\mu$ and variance $\sigma^2$, the sample mean was 10.3 and the sample variance $s^2$ was 0.13.

(a) Make a symmetric confidence interval for $\mu$ at the 99% confidence level.
(b) Test the null hypothesis $\mu = 10$ against the alternative hypothesis $\mu > 10$ at the 5% significance level.

Solution: The confidence interval is given by
$$
\hat{\mu} = \bar{X} \pm z \frac{s}{\sqrt{n}}.
$$
Here $n = 15$ and $z = F^{-1}_{t_{14}}(0.995) = 2.977$. Since $\bar{X}$ was 10.3, we get
$$
\mu = 10.3 \pm 0.28
$$
For part (b), note that since the number 10 is not in the 99% confidence interval, $H_0$ is rejected at the 1% significance level. Hence, trivially $H_0$ is also rejected at the 5% significance level.

7. (6p, only TM) Show that the Gumbel distribution is max stable, i.e. if $X$ and $Y$ are two independent Gumbel distributed random variables with the same scale parameter, then $\max(X,Y)$ is Gumbel with the same scale parameter as $X$ and $Y$. Recall that the distribution function in the Gumbel distribution is
$$
F(x) = \exp(-e^{-(x-b)/a}),
$$
where $a$ is the scale parameter.

Solution: If $X$ and $Y$ are independent and Gumbel with the same scale parameter $a$ and location parameters $b_1$ and $b_2$ respectively and $Z = \max(X,Y)$, then
$$
P(Z \leq x) = P(X \leq x)P(Y \leq x) = \exp\left(-e^{-(x-b_1)/a} - e^{-(x-b_2)/a}\right)
$$
$$
= \exp\left(-e^{-x/a}(e^{b_1/a} + e^{b_2/a})\right).
$$
Since $e^{b_1/a} + e^{b_2/a}$ is positive and the map $c \to e^{c/a}$ is continuous with image $(0, \infty)$, one can find $c$ so that $e^{b_1/a} + e^{b_2/a} = e^{c/a}$. Hence
$$
P(Z \leq x) = \exp(-e^{-(x-c)/a})
$$
for some $c$, i.e. $Z$ is Gumbel distributed.

8. (6p) Let $X_1, \ldots, X_n$ be a sample of some distribution. Show that the sample variance
$$
s^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2
$$
is an unbiased estimator of $\sigma^2 = \text{Var}(X_1)$. Show also that the sample standard deviation $s$ satisfies $\mathbb{E}[s] \leq \sigma$.

Solution: It is easy to see that $\sum_i (X_i - \bar{X})^2 = \sum_i X_i^2 - n\bar{X}^2$. We have
$$
\mathbb{E}[X_i^2] = \mu^2 + \sigma^2
$$
and
$$
\mathbb{E}[\bar{X}^2] = \mathbb{E}[\bar{X}]^2 + \text{Var}(\bar{X}) = \mu^2 + \sigma^2/n.
$$
Summing up, we get
\[ E\left[ \sum_i (X_i - \bar{X})^2 \right] = (n - 1)\sigma^2. \]

Dividing by \( n - 1 \) on both sides now gives the desired result.

For the inequality, note that
\[ 0 \leq \text{Var}(s) = E[s^2] - E[s]^2 = \sigma^2 - E[s]^2. \]

This gives \( E[s]^2 \leq \sigma^2 \). Taking square roots of both sides now gives \( E[s] \leq \sigma \).

Lycka till!
Johan Jonasson